THE PROMINENCE OF REFERRING EXPRESSIONS: MESSAGE AND LEXICAL LEVEL EFFECTS

BY

TUAN Q. LAM

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Dissertation Committee:

Associate Professor Duane G. Watson, Chair
Professor Gary S. Dell
Professor Kara D. Federmeier
Professor Cynthia L. Fisher
Assistant Professor Sarah Brown-Schmidt

Abstract

In conversation, speakers produce some words with greater intensity, longer duration, and higher fundamental frequency (F0) than other words. By making different words in a sentence more prominent than other words, a speaker can change the meaning implied by a sentence. This thesis explores the relationship between processing in the language production system and the prominence of referring expressions. In particular, this thesis focuses on the effect of processing at the message and lexical levels of language production. Across seven experiments, I examine three factors that affect prominence: predictability, repetition, and partner identity. Based upon these the results, I argue that these factors can be separated into factors that operate at the level of the message and factors that operate at the level of lexical access. Furthermore, the results suggest that message level factors and lexical level factors affect prominence in different ways: lexical level factors lead to differences in spoken duration cross-linguistically whereas message level factors manifest differently across different languages.

Authors' note

I would like to thank Allison Potter, Cara Ader, Nicole Nash, Ashley Turk, Carl Ruark, Aimee Roten, Kevin Schultz, Ilsoo Hong, and Sangeun Park for helping with data collection and transcription. I would also like to thank Jennifer Arnold for assistance in formulating the initial ideas for this project and Gary Dell and the members of the Communication and Language Lab for helpful comments on this project. This project was supported by a grant from the National Institutes of Health R01 DC008774. Additionally, the author of dissertation was supported by a training grant from National Institutes of Health T32-HD055272.

Table of Contents

Chapter 1: Introduction

Prosodic prominence

Prosodic prominence refers to the parts of an utterance that "stand out" compared to other parts of an utterance. Acoustically, these words are generally produced with greater intensity, a higher fundamental frequency (F0), or longer duration than the surrounding context. Traditionally, prosodic prominence has been described in two different ways. In one tradition, prominence is defined as a discrete linguistic construct called a pitch accent, which occurs on words that are new or focused. Pitch accents are typically marked with a change in F0, intensity, and duration, and different types of pitch accents play different roles in the discourse (e.g. Pierrehumbert, 1980; Pierrehumbert & Hirschberg, 1990). This approach is mainly concerned with where prominence is placed and the type of pitch accent that is used to convey prominence (Gussenhoven, 1983; Pierrehumbert & Hirschberg, 1990; Selkirk, 1995; Schwarzschild, 1999). ToBI, one of the more prominent discrete theories of prominence argues for two simple tones, high (H) and low (L), that combine together with lexical stress (*) to form different kinds of pitch accents (Pierrehumbert & Hirschberg, 1990). The two most commonly discussed pitch accents on this theory are the H* and L+H* accents. The H* pitch accent is marked by a rise in F0 before the most stressed syllable in a word. This type of pitch accent is believed to mark focus of a sentence because it tends to occur when a new referent enters a discourse. As a result, this type of accent is often also called the presentational stress accent. The L+H* pitch accent starts with a low F0 that rises before the stressed syllable. This pitch accent is often described as a contrastive stress marker. It tends to occur in situations where the focused word contrasts with a previously mentioned alternative.

This discrete prominence approach has its roots in the theoretical linguistics literature where researchers describe prominence in terms of the rules that govern placement of prominence. For example, Gussenhoven (1983) describes prominence in terms of a finite set of contours that are overlaid on words to convey different semantic meanings. Pierrehumbert & Hirschberg (1990) also present a semantic analysis of prominence but end up with compositional tones instead of pitch contours. However, despite these differences, these models make the same claim that prominence is discrete and finite.

The other general class of theories describes prominence in terms of its continuous acoustic-phonetic form: prominence correlates with increases in fundamental frequency (F0), duration, intensity, and intelligibility (Bard et al., 2000; Bell et al., 2003; Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Fowler & Housum, 1987; Jurafsky, Bell, Gregory, & Raymond, 2001; Watson, Arnold, & Tanenhaus, 2008). Many of these models are based on data from corpora with a wealth of data on continuous variables. One prominent proposal is Aylett & Turk's (2004) smooth signal redundancy hypothesis which tries to relate prominence to information content. This idea is based on information theory, which argues that the most efficient form of information transfer is one in which the rate of information flow is relatively constant. In information theory, the information content of a unit is measured by the predictability of that unit. Predictable units have less information while less predictable units have more information. Aylett & Turk argue that prominence acts to maintain a constant rate of signal redundancy by lengthening less predictable syllables and reducing predictable syllables. Therefore, the degree of predictability will directly affect prominence such that predictable elements will be shortened and less predictable elements will be lengthened.

One concern with the acoustic approach is that it is not clear which acoustic correlate should be measured. While, these correlates of prominence often covary; they do not perfectly co-occur (Bard et al., 2000; Kochanski, Grabe, Coleman, & Rosner, 2005; Cole, Shattuck-Hufnagel, & Mo, 2010). Whereas many researchers advocate using duration (Aylett & Turk, 2004; Jaeger, 2006; Levy & Jaeger, 2007), others have favored other metrics (e.g. Kochanski, Grabe, Coleman, & Rosner, 2005; Ladd & Morton, 1997). One solution may be to measure many different metrics. It is possible that all of these metrics matter, but for different reasons. By keeping track of many metrics one might be able to distinguish between two effects that are often correlated.

A number of factors have been shown to affect prominence, such as repetition (Fowler & Housum, 1987; Bard & Aylett, 1999, Bard et al., 2000; Aylett & Turk, 2004; Pluymaekers, Ernestus, & Baayen, 2005a; Bell et al., 2009), frequency (Gregory, Raymond, Bell, Fosler-Lussier, & Jurafsky, 1999; Fosler-Lussier & Morgan, 1999; Jurafsky et al., 2001; Pluymaekers, Ernestus, & Baayen, 2005b), and transitional probability (Bell et al., 2009; Jurafsky et al., 2001; Kidd & Jaeger, 2008). For example, Fowler and Housum (1987) found that previously mentioned words in a corpus of recorded speech are shorter and less intelligible than words that have not been previously mentioned. Similarly, in recorded speech generated from a referential communication task, Bard and Aylett (1999) found that repeated words are less intelligible to listeners than non-repeated words, and other work has shown that listeners interpret prominence as a cue to new information in on-line sentence processing (e.g. Dahan, Tanenhaus, & Chambers, 2002). Lexical frequency is also linked with prominence (Zipf, 1929). High frequency words are produced with shorter durations than low frequency words (Gregory et al., 1999; Fosler-Lussier & Morgan, 1999; Jurafsky et al., 2001; Pluymaekers, Ernestus, & Baayen, 2005b). Lexical

frequency also affects affix duration. When affixes are attached to infrequent words, the affixes are longer than when they are attached to more frequent words (Pluymaekers, Ernestus, & Baayen, 2005b). Finally, transitional probability can also affect prominence (Gregory et al., 1999; Jurafsky et al., 2001; Pluymaekers, Ernestus, & Baayen, 2005a; Kidd & Jaeger, 2008). When the transitional probability of a word is high, the acoustic realization of the word is reduced compared to when the transitional probability is low (Jurafsky et al., 2001).

Multiple sources of prominence

Prominence effects are often described as if they come from a single source (e.g. Aylett & Turk, 2004; Frank & Jaeger, 2008; Jurafsky et al., 2001; Bell et al., 2009). One of the reasons for this may be because the underlying factors that affect prominence are also correlated in natural speech (e.g. Arnold, 1998; Watson, Arnold, & Tanenhaus, 2008). For example, importance and predictability both seem to affect prominence. However, important words are generally less predictable in natural speech, so it is unclear whether the effect of predictability is simply an effect of importance or vice-versa (Watson, Arnold, & Tanenhaus, 2008).

Watson (2010) argues that prominence is best explained as the result of a combination of effects from multiple sources instead of one source. Consider a study by Watson, Arnold, & Tanenhaus (2008) in which they attempted to disentangle the correlation between importance and predictability. The task was a modified version of the game Tic-Tac-Toe. Tic-Tac-Toe is a game played on a 3x3 grid of squares where two players take turns marking the squares with the goal of placing three marks in a horizontal, vertical, or diagonal line. Tic-Tac-Toe has the advantage that the important moves are predictable and the less important moves are less predictable, which is the exact opposite pattern as in natural conversation. This meant that it was

possible to tell whether predictability or importance was the critical factor determining prominence. In the modified version, the squares were number 1-9 and instead of placing their marks themselves, participants had to name the square where they wished to place their mark. The results showed that importance leads to greater intensity whereas predictability leads to shorter duration. By tracking both intensity and duration, Watson, Arnold, & Tanenhaus (2008) were able to show that importance and predictability both affect prominence. This result is consistent with a multiple source account of prominence. Throughout this thesis, I will also take a multiple source approach when examining effects of prominence. This should allow me to examine effects that are normally correlated in natural speech.

Prominence production

Prominence effects are often described as arising from some guiding principle such as information density (e.g. Aylett & Turk, 2004) or audience design (Fowler & Housum, 1987; Fowler, 1988). For example, if a word is given in a discourse, speakers may reduce the given word because it is predictable and therefore easy for the listener to identify (Fowler & Housum, 1987; Fowler, 1988). However, it is also possible that prominence differences are the result of speaker internal production processes (Jurafsky et al., 2001; Bell et al, 2003; Bell et al., 2009). For example, repeated words may be produced with reduced prominence because of prior retrieval makes them easier and faster to retrieve later. As a result of faster retrieval, repeated words are produced with reduced prominence compared to non-repeated words.

Most models of language production agree that production has three main components: a message component, a grammatical component, and a phonetic component (Bock & Levelt, 1994; Bock, 1995; Levelt, Roelofs, & Meyer, 1999). The message component captures the

speaker's communicative intent and perspective. This component reflects what the speaker is currently thinking about and what information the speaker may wish to convey. The grammatical component is responsible for lexical selection. Lexical selection is the process by which speakers transform a thought into words. It is at this level of production that speakers select the words they will use in their utterances. Finally, the phonetic component is responsible for mapping selected words onto the sound structure for a language.

Prominence effects could arise from factors in any one of these levels; however, it is often difficult to determine whether an effect is rooted at the level of the message, lexical retrieval, or phonetic encoding. For example, predictability is correlated with prominence such that predictable words are produced with reduced prominence and less predictable words are produced with greater prominence. One possible explanation for this effect is that speakers produce less predictable words with greater prominence in order to aid their listeners in word identification. This would be a message level explanation because speakers are intentionally modulating prominence to provide additional information, in this case, the fact that the referent is unexpected for their addressees. Another possible explanation for the predictability effect is that when a referent is unexpected, the words for the appropriate referring expression have received less activation from the supporting context, and as a result, speakers may have more difficulty retrieving the appropriate word for a particular referent. This increased difficulty may lead speakers to produce the word with greater prominence. This would a be lexical level explanation because the predictability effect would be a direct consequence of difficulty selecting the correct lexical items to describe the referent. The primary goal of this dissertation will be to distinguish between prominence effects that are the result of message level factors and prominence effects that are the result of lexical level factors.

Thesis Structure

In this thesis I will explore how prosodic realizations of referring expressions vary as a product of the context in which they are spoken. First, I will demonstrate that prominence differences can arise from within the speech production system itself. From there, I will examine how different referential factors affect prosodic prominence with the goal of determining which effects are driven by differences at the message level of production and which effects originate at lower levels of production, such as lexical access or articulation.

In Chapter 2, I discuss whether the multiple source account can explain effects of repetition on prosodic prominence. A number of studies have shown that repeated words are reduced in prominence compared to non-repeated words (e.g. Fowler & Housum, 1987; Aylett & Turk, 2004; Bell et al., 2009; Galati & Brennan, 2010) and that predictable words are reduced compared to less predictable words (Aylett & Turk, 2004; Jurafsky et al., 2001; Bell et al., 2009; Watson, Arnold, & Tanenhaus, 2008). However, predictability and repetition are correlated in natural speech such that repeated words are predictable (Arnold, 1998). As a result, it is unclear whether effects of repetition are simply effects of predictability. In Chapter 2, I will discuss three experiments that are designed to test whether effects of predictability and repetition can be dissociated. The results of these experiments will lay the foundation for discussion of whether prominence effects arise from differences at the level of the message or the level of grammatical encoding.

Chapter 3 addresses the issue raised in the above sections about whether prominence differences are the result of speakers designing utterances for their listeners or the result of speaker-internal processing differences. I do this by comparing how speakers behave when speaking to different types of addressees and whether this effect could interact with predictability.

In Chapter 4, I address the question of whether repetition reduction effects are the result of differences at the level of the message or differences at the level of lexical retrieval. To do this, I will break down the effect of repetition into referent repetition and word form repetition with the intent of dissociating potentially independent contributions from both types of repetition.

Finally, in Chapter 5 I summarize the results from the experiments discussed in Chapters 2-4 and discuss their implications for a theory of how message level and lexical level factors affect prominence in production.

Chapter 2: Repetition is easy: Why repeated referents have reduced prominence[1]

Theories of acoustic prominence

While it is clear that repetition, frequency, transitional probability, and predictability in general all affect prosodic prominence, it is less clear why these effects exist. One proposal is that prominence differences are the result of speakers optimizing the acoustic signal for comprehension (Aylett & Turk, 2004; Frank & Jaeger, 2008; Lieberman, 1963; Fowler & Housum, 1987). If words are new, infrequent, and less predictable, then they may be difficult to identify in running speech. As a result, speakers may articulate these words more clearly to facilitate processing by the listener. Repeated, frequent, and predictable words are readily identifiable, so there is less need to articulate these words carefully (Fowler & Housum, 1987).

More recently, some accounts of prominence have appealed to information theoretic principles to explain differences in prominence across words. One example is Aylett & Turk's (2004) Smooth Signal Redundancy Hypothesis. According to this account, effects of repetition, frequency, and predictability on duration can be reduced to one thing: language redundancy (Aylett & Turk, 2004). Language redundancy is the predictability of a syllable, word, or syntactic structure in a linguistic context. It is modulated by a number of factors including lexical frequency, syntax, and pragmatics. In this framework, effects of repetition can also be linked to redundancy. Because speakers are more likely to refer to previously mentioned referents than new referents (Arnold, 1998), the repeated mention of a word is more predictable than the mention of a new word.

---

[1] This chapter includes material that has previously been published: Lam, T., & Watson, D. (2010). Repetition is easy: Why repeated referents have reduced prominence. *Memory and Cognition*, 38:8, 1137-1146. I would like to thank my co-author Duane Watson for his contribution to this paper.

According to the Smooth Signal Redundancy Hypothesis, speakers attempt to produce a signal in which the amount of redundancy remains relatively constant throughout production. Aylett & Turk (2004) propose that prosodic prominence's primary role is to smooth the information profile of a word. This is accomplished by reduction of syllable duration when a word is redundant: expected words are produced with shorter durations and unexpected words are produced with longer durations so that the amount of information conveyed is evenly distributed over time. Thus, the Smooth Signal Redundancy Hypothesis can explain the effects of repetition, frequency, and contextual predictability on prominence reduction.

While information theoretical accounts argue that reduction is for the benefit of the listener, this does not entail that speakers are explicitly modeling individual listeners such that specific knowledge is linked to a particular listener. Rather, speakers may be modeling a generic listener (Isaacs & Clark, 1987) whom they expect will have a certain set of beliefs or expectations. Taken to the extreme, speakers may be modeling predictability for any listener in general (Brown & Dell, 1987). Proponents of this approach argue that it is a model of communication at the computational level, rather than algorithmic level, in a Marr-like (1982) framework, and they are agnostic as to how it is implemented psychologically (e.g. Aylett & Turk, 2003; Frank & Jaeger, 2004; Jaeger, 2010). Critically, whether a specific listener is being modeled or a more generic listener is being modeled, the approach at its core depends on smoothing the signal for a listener.

An alternative view is that the link between predictability and acoustic prominence is the result of speaker internal production processes, and not a means by which the speaker facilitates processing for the listener (Bard et al., 2000; Bell et al. 2009). Most researchers agree that speech production is a multi-step process beginning with message formulation, followed by grammatical

encoding, and finishing with phonological encoding (e.g. Bock & Levelt, 1994). During message formulation, speakers formulate the semantic meaning of what they are going to say. During grammatical encoding, speakers select the appropriate lexical items to convey their messages and compute word order and include lexical items (like function words) that satisfy the constraints of the grammar of the language they speak. After grammatical encoding, speakers must encode the linguistic material into a phonological representation.

Proponents of a speaker internal account of prominence argue that prominence is linked to the amount of activation associated with a word in lexical retrieval (e.g. Bell et al., 2009). The speed at which lexical items are retrieved in the course of language production is regulated by factors like word frequency, repetition, and contextual predictability: frequent, repeated, and predictable words are retrieved more quickly than infrequent, non-repeated, and less predictable words (Griffin & Bock, 1998; Jescheniak & Levelt, 1994). Bell et al. (2009) propose that the speed of lexical retrieval is linked to articulatory planning, such that words that are retrieved quickly are articulated more quickly while words that are retrieved slowly are produced slowly. Bell et al. argue that this coordination between speed of lexical retrieval and articulatory planning is a strategy used by the production system to maximize fluent speech.

Thus, under a lexical retrieval account, repeated, predictable, and frequent words are reduced because they are retrieved quickly while non-repeated, unpredictable, and infrequent words are more prominent because they are retrieved more slowly. For example, when an entity is new to discourse, it will typically be less expected (Arnold, 1998). As a result there will be very little activation of the lexical item representing that entity and retrieval will be slower than if the word was initially more activated (Wingfield, 1968; Jescheniak & Levelt, 1994). This should result in longer articulation of the word. However, a given entity will have been

previously activated and may be less difficult to retrieve than when it was first mentioned. This could occur either because of maintenance or through the slow decay of activation (Dell, 1990). Ultimately, this would lead to reduction of the word.

Although Bell et al.'s (2009) proposal centers on lexical access as the primary determinant of word duration, in principle, facilitation of processing at other levels of production could also affect word duration. If the linguistic message is easily formulated because of its frequency, the referential context, or because it is repeated, one might expect reduction. Similarly, if the phonological form of a word is highly activated, this might lead to reducing the phonological form as well. At the heart of all of these proposals is a desire to maximize fluent speech through feed forward mechanisms from earlier stages of production to the articulatory planning stage. I return to the locus of these potential effects in the General Discussion.

Finally, a third potential account of acoustic prominence is the multiple source view (Watson, 2010). Rather than assuming that acoustic prominence has a single source, under the multi-source view, the acoustic realization of a word is the product of many factors including difficulty in speech production as well as marking information for a listener. These different factors may differentially affect varying aspects of the acoustic signal. This approach differs from that of the information theoretic and the lexical access accounts, which make specific predictions about reduction (i.e. the shortening of the duration of a word). Under the multiple source account, a word's fundamental frequency, duration, and intensity may be affected by different factors in different ways. Evidence for this view comes from both the production and comprehension literature. Watson, Arnold, & Tanenhaus (2008) found that difficult moves in games of Tic Tac Toe are produced with longer duration than moves that are easy or predictable. Moves that are important to the game, like a winning move or blocking a winning move, are

produced with greater intensity than those that are not. In comprehension, some acoustic cues are used preferentially over other cues in determining linguistic structure, suggesting that these differing cues might have differing underlying sources (Isaacs & Watson, 2010; Isaacs & Watson, 2009). For example listeners use the F0 slope over a word rather than raw duration in detecting prominence that is linked to discourse status (Isaacs & Watson, 2010). Isaacs & Watson (2009) found that intensity and not duration contribute to meta-linguistic judgments of acoustic prominence, even though both correlate with prominence in production. Thus, it is possible that acoustic prominence might be the result of both speaker internal production mechanisms and facilitating comprehension, and these factors may affect the acoustic signal in different ways.

While all three accounts predict that predictability, lexical frequency, and repetition will influence acoustic prominence, this paper focuses on repetition. This is because a lexical retrieval account and an information theoretic account can potentially make differing predictions about whether repeated words will be reduced. Under a lexical retrieval account, repetition is critical for reduction. Words that have been produced before are reduced because they are easy to repeat. In contrast, under the information theoretic approach, reduction is driven primarily by how expected a word is. One way to test whether reduction is the result of information theoretic principles or the result of processes related to lexical retrieval and planning is to test whether effects of repetition and predictability are independent. If prominence is the product of smoothing the information profile of a word for the listener, unexpected words should be lengthened, even if they have been repeated. If prominence is partly the result of speaker-internal lexical retrieval processes, speakers should reduce previously mentioned words and lengthen new words, independent of whether the word is expected or not. Residual activation stemming from the previous production should cause reduction, even if the word is unexpected.

Lastly, the multiple source account allows for the possibility that both theories may play some role in acoustic prominence, possibly in different ways.

Although I have discussed the acoustic correlates of prominence very generally, it is important to note that the information theoretic account and lexical retrieval account make predictions about duration in particular. Under information theoretic accounts, changing word duration is the means by which speakers alter the word's information profile. Similarly, under lexical retrieval theories, the difficulty of producing a word affects the word's length. In the experiments below, differences in duration will be used to adjudicate between these theories although I also measure F0 and intensity to determine whether they too are linked to repetition and predictability. The latter is critical for testing the multiple source account, as it is possible that predictability and repetition might both have effects on acoustic prominence, but they may occur along different acoustic dimensions.

Previous studies of repetition and predictability have relied primarily on corpus data (e.g. Aylett & Turk, 2004; Bell et al., 2009; Jurafsky et al. 2001). However, in natural speech, repetition and predictability are highly correlated: repeated words are more predictable than non-repeated words (Arnold, 1998). Thus, it is difficult to know whether or not effects of repetition and predictability are the result of similar cognitive processes. In Experiments 1a and 1b, I address this question by altering the correlation between predictability and repetition that exists in natural speech. In a picture description task, contexts were created in which repeating a word was unexpected, and producing a new word was expected (Experiment 1a). Contexts were also created in which repeating a word and producing a new word were equally expected, in order to elicit responses for comparison with Experiment 1a (Experiment 1b). If repeated reference causes reduction even when that target is less predictable, then this would provide support for

theories that attribute reduction to factors in lexical retrieval. If repeated, less predictable target words are produced with longer duration than non-repeated, predictable words, this would support a redundancy avoidance account: predictable words are reduced to facilitate robust communication for the listener. If both repetition and predictability play a role, this would provide evidence for the multiple source account.

Experiment 1a & 1b

Method

Participants

Sixty-three undergraduate students from the University of Illinois at Urbana-Champaign participated in this experiment to earn credit in a psychology course (32 in Experiment 1a, and 31 in Experiment 1b). All participants were native speakers of American English. Five participants had to be excluded from the analysis. One participant failed to produce the second utterance on repeated trials. Another two participants were excluded because of a recording error. The remaining two participants used pronouns on repeated mention trials, which made it impossible to compare prominence on repeated and non-repeated trials.

Materials

Participants' task was to describe events on a computer screen to a confederate. Two pictures appeared on the participant's (the director) and the confederate's (the matcher) screen for each trial. On a given trial, one of the objects would shrink and then one of the objects would flash.

The stimuli were taken from a set of twelve images from Rossion & Pourtois (2001). These images were a colorized version of images originally created by Snodgrass & Vanderwart (1980). The images were used to generate six pairs of images. Images were paired so as to avoid

semantic and phonetic relatedness. The image pairs were presented side by side in the center of the screen (see Figure 2.1). Each image pair appeared 18 times during the experiment, for a total of 108 trials. Which image appeared on the left or right was counterbalanced such that each image appeared on both sides an equal number of times. Items were randomized within sets of 12 trials such that in each set, each image pair appeared twice to counterbalance the image location.

Repetition of one of the objects in the task and the likelihood of it being mentioned were both manipulated. Repetition was manipulated by varying whether the same object engaged in a shrinking and flashing event. On repeated mention trials, one of the images shrank and then the same image flashed. On non-repeated mention trials, one of the images shrank and then the other image flashed. An example of a trial from the non-repeated condition is shown in Figure 2.1.

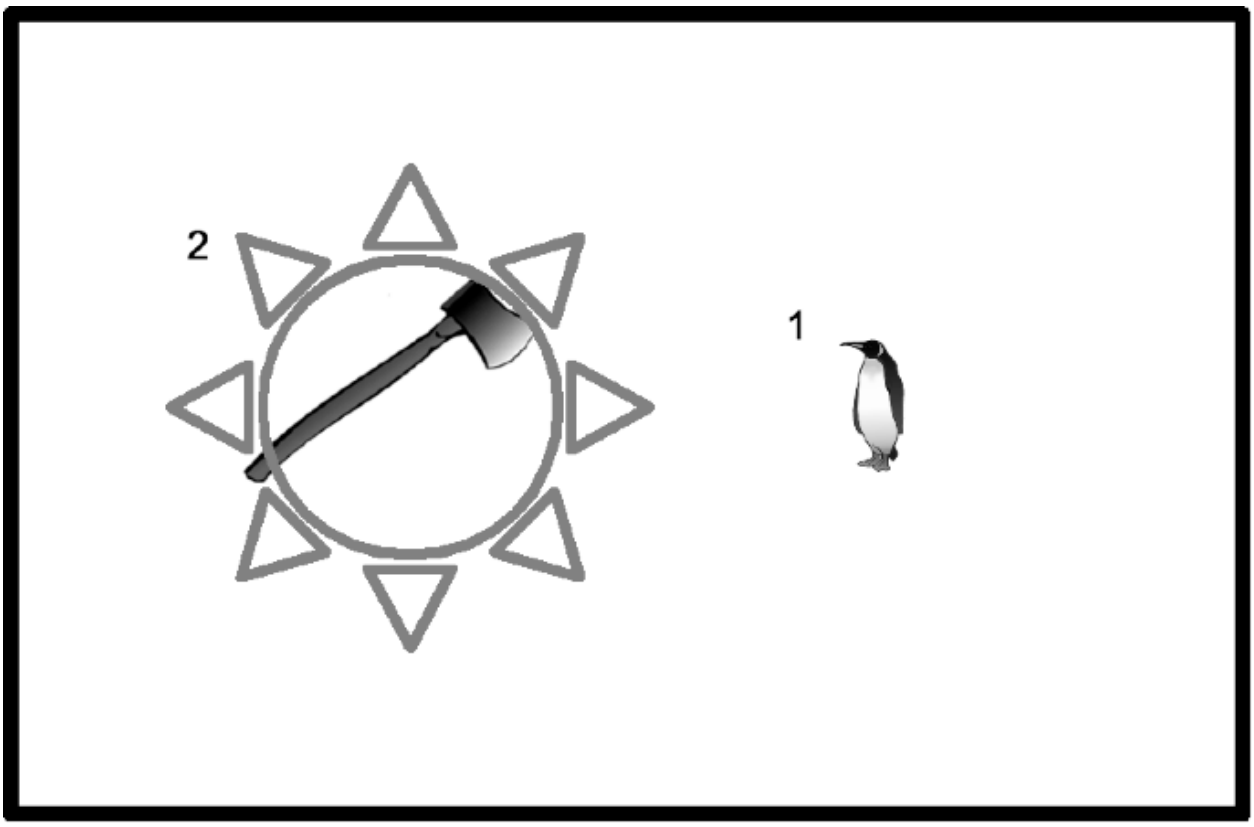

Figure 2.1: An illustration of a typical trial from Experiment 1 in the non-repeated condition.

Note: The numbers indicates the order of events. The sun indicates the flashing event.

Typical utterances are presented in (1) and (2).

(1) Repeated noun

The axe is shrinking…The axe is flashing.

(2)Non-repeated noun

The penguin is shrinking…The axe is flashing.

In order to manipulate predictability, a training block, followed by a test block, was used to manipulate speaker expectations about what object on the screen would flash. The training block consisted of 96 trials and the test block consisted of 12 trials. In the training block of Experiment 1a, repeated mention trials were much less predictable than non-repeated mention trials. In the training block, only six of the trials were repeated mention trials. The order of the conditions was pseudo-randomly permuted such that no repeated mention trial occurred within five trials of another repeated mention trial. Moreover, the first eight trials were all non-repeated mention trials. In the training block of Experiment 1b, repeated mention trials and non-repeated mention trials were equally likely and the order of the conditions was randomly permuted. Each training set was randomized such that every participant was presented with a different item order during the training block.

The purpose of the test block was to determine what effects repetition and the expectations established in the training block had on speaker productions. The test block consisted of 12 trials. The transition between training and test block was not marked, so participants were unaware of the transition. During the test block participants were exposed to each pair of images twice, one time in the repeated condition and one time in the non-repeated condition. As with the training block, the order of items during the test block was randomized. No pair of items appeared twice in a row during the test block. In Experiment 1a, there were 30 unique lists such that each participant was exposed to a unique order of trials during the test block. The order in which items were presented in the test block was matched across Experiments 1a and 1b.

The experiment was programmed using MATLAB with the Psychophysics toolbox version 2.54 installed. Participant utterances were recorded at a frequency of 44kHz.

Procedure

Before the experiment, participants were shown a video in which two research assistants were presented completing the task. This video was used to both instruct the participant on how to complete the task and to prime the participant with the construction, "the Noun 1 is shrinking…the Noun 2 is flashing." Then participants were told that they would play the role of the director while the research assistant would play the role of the matcher. The director sat at a computer facing away from the matcher's computer. The pair completed six practice trials before beginning the actual experiment. The images in the practice trials were the same images as those used in the experimental trials, and each pair was used only once. All practice trials were non-repeated mention trials.

At the beginning of each trial, two images appeared. After 500ms, one of the images shrank. Two seconds after the shrinking event, either the same image flashed, or the other image flashed. The participant was instructed to describe each event as soon as he knew what was happening. The matcher, meanwhile, clicked on one of four buttons to make her screen match the director's screen. These buttons corresponded to the shrinking event and flashing event for each image. The trial ended when the matcher notified the director that she was finished matching the second event. Occasionally, the speaker accidently misnamed an object. On those trials, the matcher provided feedback that she did not have the object thereby prompting the director to correct his utterance. Otherwise there was no explicit feedback aside from confirmation of completion of the trial. The pair completed 108 trials of which the last 12 were recorded and labeled using Praat, a speech analysis program developed by Boersma & Weenink

(2007). The target words during the first and second utterances were analyzed for mean F0 over the word, the maximum F0 excursion over the word, the minimum F0 over the word, word duration, and mean intensity. I also computed the proportion of the total utterance duration due to the target word in each condition (hereafter called target proportion).

Results

The data were analyzed using a linear mixed effects regression model with subject and item as random intercepts and slopes using the lmer function in the lme4 package in R (Baayen, 2008). Like ANOVA, this method accounts for the variance due to subjects and items; however, this method can account for variance of multiple random factors simultaneously (Baayen, 2008). Model comparisons were conducted using likelihood a ratio test to find the best fit random slopes and intercepts models. Random slopes did not significantly increase model fit for any reported model and are therefore not reported. Reported p-values were obtained from Markov chain Monte Carlo (MCMC) sampling using the language R package (Baayen, 2008). The production of the second target word across conditions was compared. I also compared the production of the first target word in a trial (Noun 1) to the production of the second target word (Noun 2) for trials in the repeated condition. All predictor variables were contrast coded, and as a result of the balanced design, the predictors were centered on the mean.

In Experiment 1a, there were effects of repetition but no effects of predictability. Non-repeated Noun 2's were produced with greater duration ($t=3.421$, $p<0.001$, $\beta=17.0$ S.E.$=4.96$) than repeated Noun 2's; however, there was no significant difference for intensity ($t<1$). There were also no significant differences in F0 across conditions for Noun 2's. The target proportion in the non-repeated condition was also significantly greater than the target proportion in the repeated condition ($t= 2.969$, $p<0.001$, $\beta=0.0112$ S.E.$=0.00377$). The means for Noun 2 across

repetition conditions are presented in Table 2.1. In repeated trials, Noun 1 was produced with greater duration ($t=3.78$, $p<0.001$, ß=20.9 S.E.=5.54) and intensity ($t=4.98$, $p<0.0001$, ß=1.08 S.E.=0.216) than Noun 2. Noun 1 was also produced with higher maximum F0 ($t=2.98$, $p<0.01$, ß=13.2 S.E.=4.434), higher minimum F0 ($t=4.91$, $p<0.001$, ß=18.7 S.E.=3.82) and higher mean F0 ($t=6.13$, $p<0.001$, ß=14.6 S.E.=2.38) than Noun 2. The means for Noun 1 and Noun 2 in the repeated condition are presented in Table 2.2.

In Experiment 1b, non-repeated Noun 2's were produced with greater duration ($t=3.62$, $p<0.001$, ß=16.9 S.E.=4.68) and intensity ($t=2.83$, $p<0.01$, ß=0.645 S.E.=0.228) than repeated noun 2's. The target proportion was also significantly longer in the non-repeated condition ($t=3.89$, $p<0.0001$, ß=0.0139 S.E.=0.00356). As in Experiment 1a, there were no significant differences in F0 across conditions for Noun 2. In repeated trials, Noun 1 was produced with greater duration ($t=2.758$, $p<0.01$, ß=12.7 S.E.=4.62) and intensity ($t=5.06$, $p<0.0001$, ß=1.52 S.E.=0.231) than Noun 2. Noun 1 was also produced with a higher average F0 ($t=4.07$, $p<0.05$, ß=10.7 S.E.=2.64) and a higher minimum F0 ($t=3.05$, $p=0.01$, ß=10.8 S.E.=3.53). Maximum F0 did not differ significantly across conditions.

Table 2.1

Experiment 1a and Experiment 1b Noun 2 Summary

	Experiment 1a				Experiment1b			
	Non-repeated		Repeated		Non-repeated		Repeated	
Metric	M	SE	M	SE	M	SE	M	SE
Duration (ms)	393	(13.8)	376	(14.3)	356	(10.3)	339	(10.7)
Proportion	.424	(.008)	.413	(.008)	.420	(.007)	.406	(.007)
Intensity (db)	78.7	(1.14)	78.6	(1.08)	77.6	(0.97)	77.0	(1.00)
Average F0 (Hz)	167	(9.02)	165	(9.14)	156	(8.94)	155	(8.19)
F0 Maximum (Hz)	196	(11.3)	194	(11.2)	179	(10.7)	181	(10.7)
F0 Minimum (Hz)	141	(8.07)	140	(8.23)	136	(7.36)	135	(7.19)

Values in parentheses represent standard errors of the means. Note: In Experiment 1a, repeated nouns have low predictability (6.25% of trials) while non-repeated nouns have high predictability (93.75% of trials). In Experiment 1b, repeated and non-repeated nouns are equally likely.

Table 2.2

Experiment 1a and 1b. The means of Noun 1 and Noun 2 in the repeated condition

	Experiment 1a				Experiment1b			
Metric	Noun 1		Noun 2		Noun 1		Noun 2	
	M	SE	M	SE	M	SE	M	SE
Duration (ms)	397	(14.0)	376	(13.8)	352	(10.6)	339	(10.7)
Intensity (db)	79.7	(1.10)	78.6	(1.08)	78.5	(0.953)	77.0	(1.00)
Average F0 (Hz)	180	(10.1)	165	(9.14)	165	(9.60)	155	(8.19)
F0 Maximum (Hz)	207	(12.0)	194	(11.2)	188	(10.7)	181	(10.7)
F0 Minimum (Hz)	159	(8.98)	140	(8.23)	145	(8.25)	135	(7.19)

Values in parentheses represent standard errors of the means. Note: In Experiment 1a, repeated Noun 2s have low predictability (6.25% of trials). In Experiment 1b, repeated Noun 2s have relatively higher predictability (50% of trials).

Discussion

In both Experiment 1a and 1b, repeated nouns were less prominent than non-repeated nouns, providing support for a lexical retrieval account of prosodic prominence. According to these accounts, repeated words should be reduced because they have been previously activated

and are therefore easier to retrieve for a subsequent production. This was also true of the overall utterance duration and the target proportion. Surprisingly, there were no differences between Noun 2's in F0. In repeated conditions, F0 was higher for the first noun than second noun though this may have been due to the declination in pitch that typically occurs over a set of related utterances. It is possible that the descriptive nature of the task led participants to vary their pitch less than they would have in a more interactive setting.

These results are less consistent with information theoretic approaches. According to information theory based accounts, the predictability manipulation in Experiment 1a should have led to reduced duration of non-repeated, expected nouns and lengthening of repeated, unexpected nouns. In fact, the reverse occurred.

Note however that although the duration results support the theory that prominence is a result of lexical retrieval production processes, there were some differences in intensity across experiments. A post-hoc test for a condition by experiment interaction was conducted on the Noun 2 intensity data, which yielded a marginally significant interaction ($t=1.77$, $p=0.07$, $ß=0.561$ S.E.$= 0.317$). There were no differences in intensity in Experiment 1a between the repeated, unexpected condition and the non-repeated, expected condition, but in Experiment 1b, non-repeated noun 2's were produced with greater intensity than repeated noun 2's. This pattern of results suggests that the lack of intensity differences observed in Experiment 1a may have been due to the predictability manipulation. This marginal interaction suggests that predictability and repetition may in fact be two separate factors that influence the production of prominence, a possibility that is most consistent with the multiple source view of prominence outlined in the introduction.

Predictability and repetition might both have had effects on intensity, but effects of predictability were not detectable because predictability and repetition were not independently manipulated. In the contexts of these experiments, it is difficult to determine whether this occurred, because in both instances, predictability and repetition were negatively correlated. I address this issue in Experiment 2.

Experiment 2

The goal of Experiment 2 was to test whether or not predictability and repetition have separate, independent effects on the acoustic realization of a word. A shortcoming of Experiment 1a was that predictability and repetition were negatively correlated, so the effects of one factor might have obscured effects of the other.

In Experiment 2, I altered the task used in Experiment 1 so that predictability and repetition could be independently manipulated. Participants were presented with an array of twelve images. As in Experiment 1, one image shrank and then another image flashed. However, the second event was preceded by a probabilistic cue as to which object would flash. In 92% of trials, a circle appeared around the image that flashed. At the beginning of the experiment, participants were told that the circle would usually but not always indicate which object would flash. This made it possible to independently manipulate whether a given word was repeated across events and whether it was expected by the speaker. An information theoretic account predicts that expected words will be produced with shorter durations than unexpected words. This account also predicts that there should be no differences between repeated and non-repeated words because the reliability of the cue was the same in repeated mention and non-repeated mention trials. In contrast, a lexical retrieval account predicts that repetition, rather than

predictability, should affect reduction. This account predicts that repeated words should be produced with less prominence than non-repeated words. Finally, the multiple source account predicts that both repetition and predictability affect prominence and that these effects may be realized in different ways acoustically.

Method

Participants

Forty-five undergraduate students from the University of Illinois at Urbana-Champaign participated in this study in exchange for course credit. All participants were native speakers of American English. Four participants were excluded due to recording errors. One participant was excluded for failing to follow the instructions.

Materials

As in Experiment 1a and 1b, images were taken from a set of colored images by Rossion & Pourtois (2001). Ninety-six images from this set were used, of which 72 images were targets in critical trials. For each trial, a 3x4 array of images was displayed on a computer screen using MATLAB with the Psychophysics toolbox version 3.0 installed.

As in Experiments 1a and 1b, there were two events on each trial. On every trial, one of the images shrank and one of the images flashed. Predictability and repetition were manipulated in a 2 x 2 factorial design. Predictability was manipulated by circling a potential target for the second utterance immediately after the shrinking event but before the flashing event. On predictable trials, the circled image flashed. On unpredictable trials, the circled image did not flash. Repetition of the target word was also manipulated: either the same object shrank and flashed or different objects shrank and flashed. Predictability and repetition were crossed in a

2x2 design yielding four conditions: repeated-expected, repeated-unexpected, non-repeated-expected, and non-repeated-unexpected.

At the beginning of a trial, 12 images appeared on the screen. After one second, one of the images shrank. Then, after one second, one of the images was circled. The circle remained on the screen for 500 ms and then disappeared. After another 500ms, one of the images flashed. The images that shrank and/or were circled depended upon the condition for the trial. In repeated mention trials, the image that shrank was also the image that flashed. In non-repeated mention trials, the image that shrank was not the image that flashed. In expected trials, the image that was circled was the image that eventually flashed. In unexpected trials, the image that was circled was not the image that eventually flashed. An illustration of a trial from the non-repeated-unexpected condition is shown in Figure 2.2.

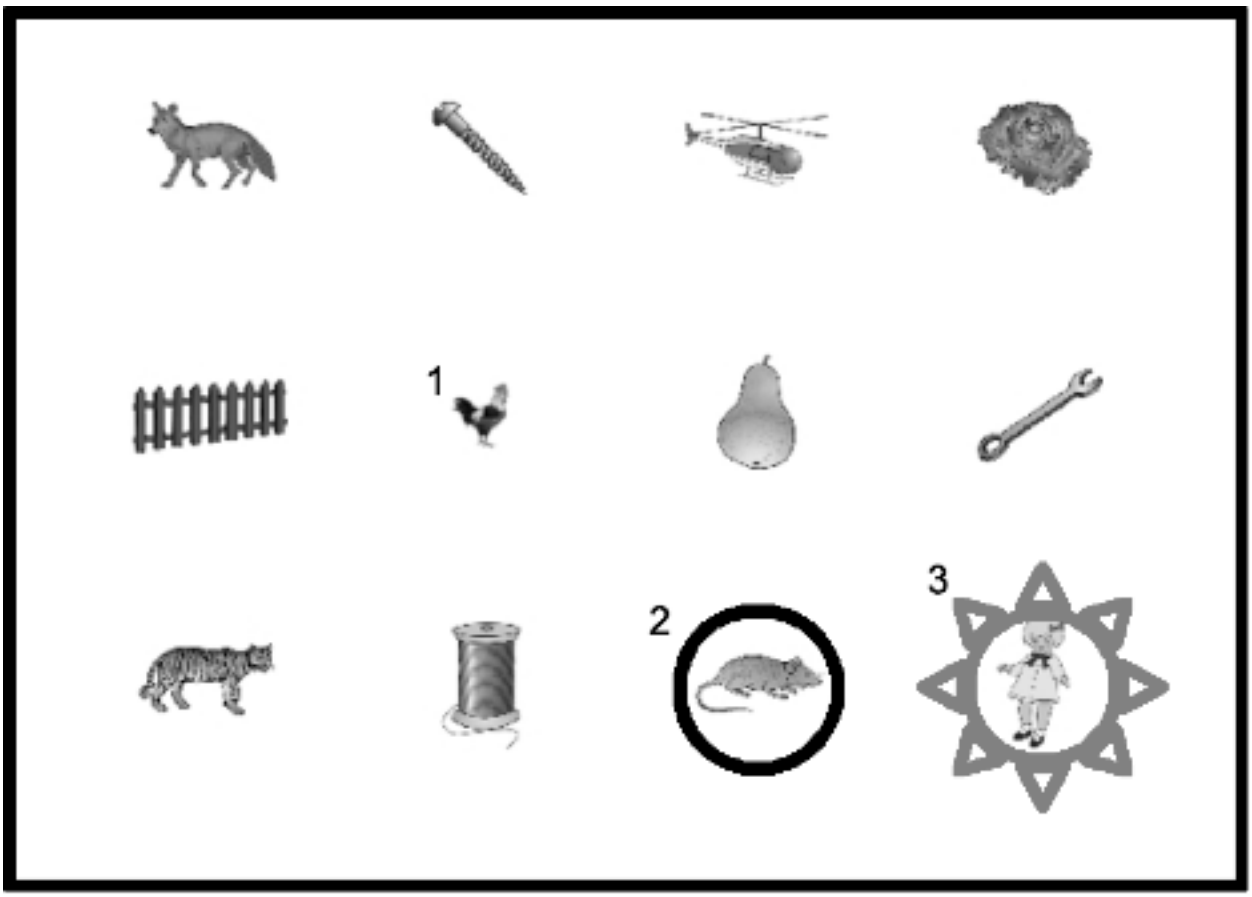

Figure 2.2: An example of a typical trial in Experiment 2 from the non-repeated expected condition. Note: The numbers indicate the order of events in the trial. The sun indicates the flashing event.

There were six critical trials for each condition for a total of 24 critical trials. There were 120 filler trials, which were all trials in which the targets were non-repeated and predictable. These filler trials were used to reinforce the predictability manipulation. Overall, the cue was reliable on 132 of 144 trials or roughly 92% of the time. The order of trials was pseudo-randomly permuted such that no two critical trials appeared in succession. On critical trials, the shrinking image, the circled image, and the flashing image were all novel as targets but may have appeared previously as filler images. After the critical trial in which these images were used as targets,

they could appear again later as targets in filler trials. Because of the potential for order effects, two pseudo-randomized target lists were used. In each list, critical items were counterbalanced using a Latin square, resulting in eight lists in total. The target on each trial was counterbalanced so that it appeared in all 12 locations an equal number of times in both critical trials and filler trials.

Procedure

Before beginning the experiment, participants were shown a video that demonstrated the task. This video was used to prime participants to use the construction "the Noun 1 is shrinking … the Noun 2 is flashing." The video was also used to inform the participants about the probabilistic cue to the flashing event. Participants were told that the circled image was frequently the image that flashed. After watching the video, participants completed eight practice trials. Five of the trials were generated from the non-repeated-predictable condition. The other three trials were from the remaining conditions, repeated-expected, repeated-unexpected, and non-repeated-unexpected. These last three conditions were included so that participants would not be surprised when they encountered them in the actual experiment.

Following the practice trials, participants immediately began the experiment. Unlike in Experiments 1a and 1b in which the participant addressed a matcher, in Experiment 2, the speaker was alone in the room and progress through the experiment was self-paced. At the beginning of a trial, one of the images shrank. The participant then described this event. Then one of the images was circled. Finally, one of the images flashed and the participant described this event. After describing the flashing event, the participant pressed a key to begin the next trial. Participants completed 144 continuous trials of which 24 critical trials were recorded. The target

word was the production of the flashing noun. The mean F0, max F0, min F0, intensity, and duration of target words were measured. I also computed the proportion of the total duration of the utterance occupied by the target word in each condition.

Results

Two targets (screw & refrigerator) were removed from analysis due to inconsistency in naming them across subjects. Of the remaining targets, 47 trials were removed from analysis due to errors in naming the targets. This led to a loss of 5.34% of the total trials.

Means are presented in Table 2.3. The data were analyzed using linear mixed effects regression with subject and item as random effects. All predictor variables were contrast coded and centered. As in Experiment 1a and 1b, model comparisons were conducted using a likelihood ratio test to find the best fit random slopes and intercepts models. Again, random slopes did not significantly increase model fit for any reported model and are therefore not reported. All reported models include random intercepts.

Both repetition and expectedness were reliable predictors of intensity, and both factors together were better predictors of intensity than either factor alone. Non-repeated words had greater intensity than repeated words ($t=2.73$, $p<0.01$, $ß=0.4802$ S.E.$=0.176$), and unexpected words had greater intensity than expected words ($t=3.77$, $p<0.001$, $ß=0.6617$ S.E.$=0.176$). The relative size of the regression coefficients suggests that predictability had a larger effect than repetition on intensity.

Only repetition was a reliable predictor of raw noun duration. Non-repeated words were longer than repeated words ($t=6.58$, $p<0.0001$, $ß=30.0$ S.E.$=4.55$). Predictability did not reliably predict raw duration.

There were effects of both repetition and predictability on target proportion. Non-repeated words had a greater target proportion than repeated words ($t=5.37$, $p<0.0001$, ß=0.0145 S.E.=0.00270) and unexpected words had a greater target proportion than expected nouns ($t=2.70$, $p<0.01$, ß=0.0073 S.E.=0.00270). For target proportion, the relative sizes of the regression coefficients suggest that repetition has a larger effect than predictability.

For measures of maximum F0 and average F0, there was a significant interaction between repetition and expectedness ($t=2.954$, $p<0.01$, ß=25.1 S.E.=8.51; $t=2.362$, $p<0.05$, ß=8.57 S.E.=3.65 respectively). Post-hoc paired t-tests revealed that non-repeated nouns had a higher average F0 than repeated nouns in the unexpected condition ($t_{39}=3.367$, $p<0.001$); expected nouns showed no difference across repetition conditions ($t_{39}=0.088$, $p=0.93$).

Table 2.3

Experiment 2 Noun 2 Summary

	Non-repeated				Repeated			
Metric	Expected		Unexpected		Expected		Unexpected	
	M	SE	M	SE	M	SE	M	SE
Duration (ms)	446	(13.5)	458	(11.9)	421	(10.9)	428	(12.1)
Noun Proportion	.401	(.006)	.420	(.006)	.398	(.005)	.404	(.006)
Intensity (dB)	58.4	(0.81)	59.2	(0.88)	58.2	(0.82)	58.7	(0.91)
Average F0 (Hz)	170	(7.49)	177	(7.62)	173	(6.84)	168	(6.70)
F0 Maximum (Hz)	206	(9.28)	220	(9.54)	213	(10.0)	198	(8.87)
F0 Minimum (Hz)	146	(6.51)	149	(6.71)	145	(5.84)	144	(6.07)

Values in parentheses represent standard errors of the means.

Discussion

The results from Experiment 2 suggest that both repetition and predictability play independent roles in the production of prosodic prominence. Both factors affected the duration and intensity of the target word. However, predictability was the stronger predictor of intensity while

repetition was the stronger predictor of duration. The data from Experiment 2 suggests that the effects of repetition obscured effects of predictability in Experiment 1a, particularly for intensity. Recall that in Experiment 1a, these two factors were placed in opposition to one another, making it difficult to know whether each factor contributed independently to the acoustics of the target word. The weak effect of predictability on duration is evidence against a strong redundancy account of prominence because these theories predict that duration specifically should be reduced. This is not due to a weak effect of predictability overall as the predictability effect was relatively strong for intensity. This seems to suggest that these predictability and repetition effects on prominence come from different sources. This would be most consistent with the multiple source account.

Interestingly, there was a reliable interaction between predictability and repetition with respect to F0 such that unexpected words were produced with higher F0 in the non-repeated condition than in the repeated condition, but expected words were produced with similar F0 in the repeated conditions. This suggests that the lack of an effect in Experiment 1 was not due to the nature of the task. This interaction was not predicted by any of the theories discussed above, and suggests that the presence of both factors may be necessary for triggering a higher F0. Future work will need to investigate why effects on F0 appear to be qualitatively different than effects on duration and intensity.

One potential concern is that because there was no overt listener (unlike Experiment 1), these data may not be useful in evaluating information theoretic approaches. There are two things to note. The first is that, as discussed in the introduction, in the information theoretic frameworks that have been proposed, it is not critical that a specific listener be present. Frank & Jaeger (2008) are agnostic as to whether a specific listener is modeled, and argue that these

information theoretic principles apply to communication more generally. The second is that despite the absence of a listener, there were clear effects of predictability. The manipulation of predictability was clearly strong enough to elicit differences. However, these differences appeared primarily in measures of intensity, rather than duration.

If repetition and predictability effects on prominence arise from different sources, what are the sources? For the repetition effect, the current results are consistent with a lexical access based account of prominence reduction. According this account, repeated words are reduced in duration because previous retrieval makes them easier to retrieve subsequently. If repetition reduction is due to activation from prior lexical access, then effects of duration reduction should apply in languages that have different prominence patterns from English. This is because what is critical is that prior lexical retrieval leads to faster subsequent retrieval. As a result of faster lexical retrieval, words are produced with shorter durations. As for the predictability effect, it is possible to revise the redundancy account such that redundancy leads to reduction of prominence in general, including both duration and intensity. This account would allow speakers to use either dimension to convey prominence and would still predict that words that are predictable would be produced with lower prominence than words that are less predictable. Such an account would allow for languages to vary in which acoustic correlate of prominence they treat as important.

Experiment 3

The goal of Experiment 3 is to test whether the predictability and repetition effects manifest in the same way in languages other than English. If the repetition effect is due to low level priming, such as activation from prior lexical access, then effects of duration reduction should apply in languages that have may have different prominence patterns from English. In this experiment I

will examine how the predictability and repetition effects manifest in Korean, a language that differs from English in how it expresses prominences and stress.

The stress pattern in English has strong and weak syllables with “pitch accents” typically occurring on nuclear strong syllables. These strong syllables in English are typically marked with tense vowels while weak syllables are typically marked by used lax vowels. However, even when words carry the same vowels, there can be differences in weak and strong syllables. These strong syllables are produced with greater duration and higher intensity than weak syllable. Moreover, there is evidence that in English, prominence is typically carried on the vowel of the strong syllables. However, Korean does not carry the strong weak distinction that English carries, and therefore, Korean cannot have pitch accents. Rather, stress is carried in a supra segmental accentual phrase that typically ends in a high tone, which is a difference in F0 (Jun, 1993; 1998). Therefore, it is possible that whereas prominence is most reliably shown in duration, intensity, and F0 with English (e.g. Kochanski et al., 2005; Cole, Shattuck-Hufnagel, & Mo, 2009), it may appear primarily on F0 for Korean. Moreover, if the hypothesis that durational differences reflect priming in the production system is correct, there should still be an effect of repeated mention on duration, but there may not be an effect on intensity. This is because in a lexical access theory the mechanisms that lead to repetition reduction in duration should be universal across languages. Prior activation of a lexical item should lead to easier repeated access, which in turn can lead to reduced duration (Bell et al., 2009).

Method

Participants

Twenty-five speakers of Korean participated in return for $20 or for credit in an introductory course. All participants were native speakers of Korean with English as a second language. All participants started learning Korean from birth and lived in a Korea during primary school with the majority of participants living in Korea until they began university studies. Of the twenty-five subjects, twenty-two were native speakers of the standard Seoul-Gyeonggi dialect of Korean while the remaining participants were native speakers of Gyeongsang dialect who were familiar with Seoul dialect.

Materials

The materials were similar to those used in Experiment 2 from Chapter 2, except that this experiment was conducted in Korean instead of English. Again, subjects completed an image description task. However, the exact images used differed somewhat from the images that were used in the English version to control for properties of name agreement in Korean. In order to chose the images, I first asked three native speakers of Korean to provide Korean labels for the 260 images from Rossion & Pourtois (2001). Of these 260 images, 193 of the images had consistent labels across the three labelers. Next, I removed images that had names that seemed to be loan words from English. This left 150 images.

Because Korean is a case marking language, I also took into account the particle used for each word. Korean has two variants of the subject marker (ㅇ] ~ [i] and 가 ~ [gɑ]). If the word ends with a coda consonant, then the particle is ㅇ]. If the word ends with a vowel then the particle

is 가. For consistency, all critical target words ended with a vowel, so speakers used the 가 particle. Filler targets used a mix of 이 particle words and 가 particle words.

Procedure

The procedure for this experiment was identical to the English version, except that it was conducted in Korean. Participants first watched an example video of a participant completing a few trials. Then participants were then presented with instructions for how to complete the task and given eight practice trials. During the practice trials the experimenter gave the participant feedback in Korean. Following the practice trials, speakers completed 144 experimental trials.

Predictions

The most important acoustic metric for this study is duration. If the repetition effect is due to priming in the production system leading to reduction in duration, then Korean should show reduced duration for repeated mentions. Moreover, repetition should show a stronger effect for duration than for intensity. If Korean does not show a repeated mention effect on duration, then it suggests that the repeated mention effect on duration is not due to simple priming in the production system, but may be due to other factors.

The predictions for predictability are less clear. English shows an effect of predictability on intensity: predictable words are produced with lower intensity than less predictable words. , English also shows an effect of predictability on duration: predictable words are produced with shorter durations than less predictable words. However, this may be due to the fact that prominence in English is typically carried by stressed syllables. Because Korean does not have this strong weak contrast, the effect of predictability may not appear on intensity. The most

likely acoustic metric for predictability to appear on, if not intensity, would be F0. This is because prominence in the Korean accentual phrase correlates with F0 changes.

Results

The data were analyzed using multilevel linear mixed effects regression. For each acoustic metric, I tested for random slopes and intercepts for subject and target word. I used AIC to determine the model with the best random effects structure for each acoustic metric. For each acoustic metric collected, the model with the best random effects structure included random intercepts for subject and target word but no random slopes. I will report only the models with the best fitting random effects structure. Reported p-values are from Markov chain Monte Carlo (MCMC) sampling.

The analysis for target duration was conducted on the log of the duration values. The patterns did not change when analysis was conducted on raw durations. For target duration there was a main effect of repetition and predictability (see Figure 2.3). Repeated words were produced with shorter duration than non-repeated words ($t=4.83$; $p<0.0001$). This pattern is consistent with the lexical access account of repetition reduction because it predicts that repetition reduction is due to faster access of previously mentioned words. However, predictable words were produced with *longer* duration than less predictable words ($t=-2.93$; $p<0.01$). This pattern is the opposite of the pattern for English in which predictable words were produced with *shorter* durations than less predictable words and is also evidence against the redundancy account of prominence. There was no significant interaction of repetition and predictability ($t<1$). The pattern for target proportion (i.e. target duration/duration of utterance) was similar. Repeated words had a smaller target proportion than non-repeated words ($t=3.98$, $p<0.001$) and predictable

words had a larger target proportion than non-repeated words (t=-2.42, p<0.01). As with raw duration, there was no significant interaction between repetition and predictability (t<1).

For F0, I analyzed the log of average F0, maximum F0, minimum F0, and the range of F0 (maximum F0– minimum F0). There were no significant differences for average F0 and F0 min so I will only report the analysis for maximum F0 and the range of F0. For maximum F0 there was a significant effect of repetition (see Figure 2.4). Repeated words were produced with lower maximum F0 than non-repeated words (t=2.94, p<0.01). There was no significant effect of predictability and no significant interaction for maximum F0. The pattern for range of F0 is similar to maximum F0. For range of F0, there was also a significant main effect of repetition; repeated words were produced with a smaller range in F0 compared to non-repeated words (t=2.15, p<0.05).

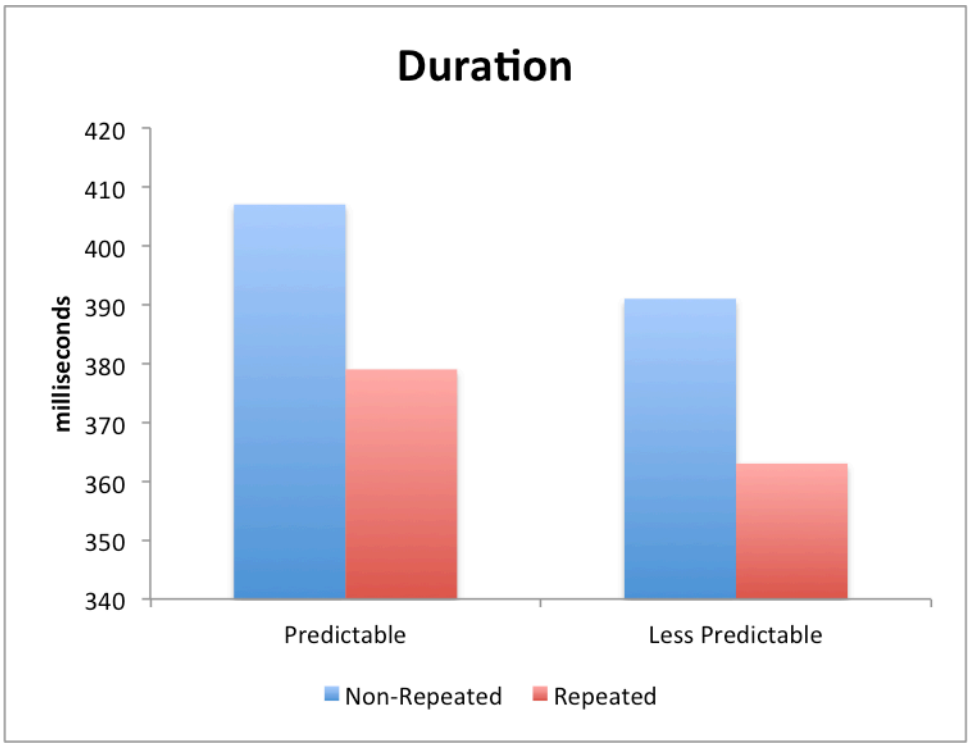

Figure 2.3: Average duration by condition for the Korean data

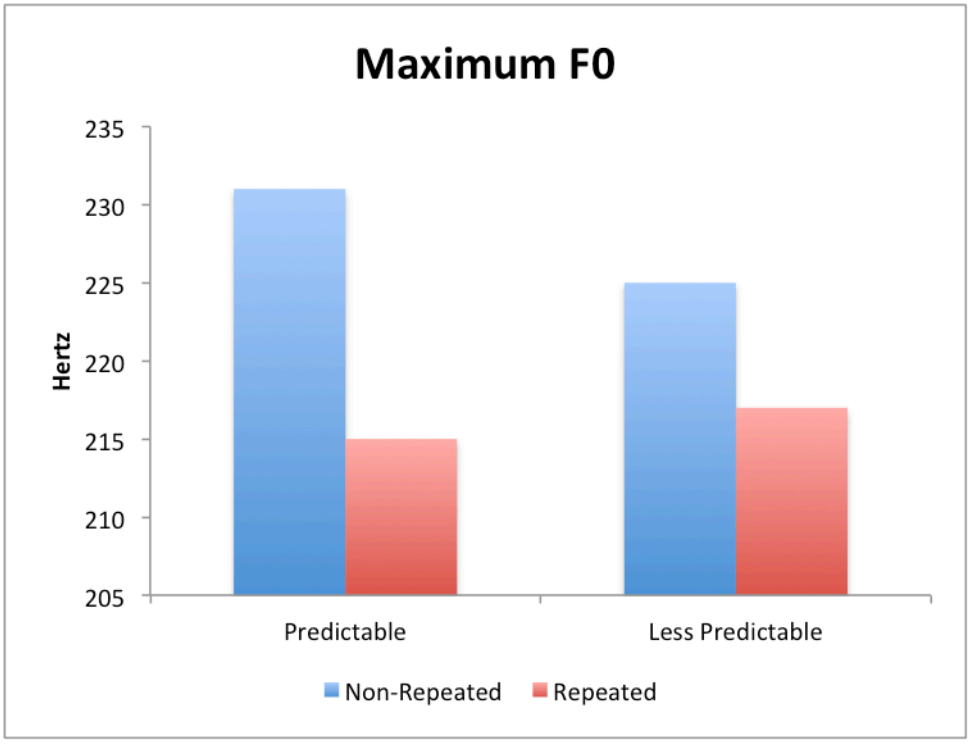

Figure 2.4: Average maximum F0 by condition for the Korean data.

Discussion

The prominence pattern in Korean again supports the idea that repetition and predictability are two separate factors that affect prominence production. The duration pattern provides evidence that repetition reduction on duration may be a language independent effect. As in English, repeated words in Korean were produced with shorter duration than non-repeated words. This provides support for the lexical access theory of prominence production which predicts that prior mention should lead to shorter duration because phoneme retrieval for previously mentioned words will be faster.

The duration results also call into question the theory that predictable elements are always produced with reduced prominence. The pattern for predictability was in the opposite direction as in English. Predictable words were produced with shorter duration than less

predictable words. This was true whether duration was analyzed as raw durations of the target noun or as a ratio of the target noun duration compared to the entire utterance. This pattern provides strong evidence against a redundancy account of repetition reduction because repeated words were produced with *shorter* durations while predictable words were produced with *longer* durations. This pattern also provides evidence against a redundancy account of prominence more generally. Recall that the redundancy account argues that speakers reduce prominence for predictable words to control the rate of information transfer such that each element has a similar amount of information density (i.e. information per unit time). This is argued to provide the mathematically most efficient rate of information transfer (Aylett & Turk, 2004). According to these accounts, predictable words have less information content and less predictable words have more information (Shannon, 1951). Using this information density metric, the Korean data actually becomes less uniform because predictable words are produced with longer durations leading to very low information density while less predictable words are produced with shorter durations leading to very high information density.

If predictability's effect on prominence is not about information density, then what might be the cause of the predictability pattern for Korean? One possibility is that the duration reduction still provides a cue to predictability for the listener. A word that is reduced compared to the typical or predicted duration still provides a binary cue. It is possible that reduction is simply how Korean signals when a word is unexpected. Another possibility is that duration works in conjunction with intensity or F0 to signal predictability via slope or spikiness. For example, if two words have the same intensity, but one word is shorter, then the shorter word has a larger intensity by duration slope. This slope could be the signal that Korean speakers use. In the Korean data, predictable words were indeed produced with lower intensity slope than less

predictable words; however, non-repeated words were also produced with lower slope than repeated words. This would create a new problem because the repetition effect would now be in the opposite direction as before. Finally, Korean and English have many phonetic and syntactic differences. One of these differences may be the cause of the difference in how Korean and English signal predictability. Regardless of the cause, the fact that Korean and English show different prominence patterns for predictability suggests that predictability is signaled differently across languages. In order to discover what this difference may be, it is important to examine prominence across a number of other languages.

The F0 data also show an interesting pattern. Repetition led speakers to produce words with lower maximum F0, and as a result, lower average F0, and a reduced range of F0. Predictability did not show any reliable effects on F0. That fact that repetition led to reduced maximum F0 in Korean is perhaps not surprising given that prominence in Korean is typically carried on the accentual phrase. The accentual phase in Korean is correlated with increases to duration and pitch expansion. It is possible that the F0 effect may have been a byproduct of Korean speakers producing a pitch accent for the Korean accentual phrase; however, in order for this to be true, duration differences must arise from sources other than just pitch accenting alone If the F0 effect and duration effects are both by-products of Korean speakers producing a pitch accent, duration and F0 should have the exact same patterns. Rather, this pattern suggests that effects on F0 and duration in Korean are sensitive to different factors. This would be compatible with the multiple source account of prominence (Watson, 2010).

Finally, while the Korean results provide some support for the lexical access account of prominence, it important to try to replicate this effect in other languages. Much of the work investigating the relationship between redundancy and prominence has focused on English.

However, English has specific properties that may cause it to behave differently from other languages such as lexical stress or SVO word order. In fact, the duration pattern for predictability in Korean demonstrates this quite clearly. The fact that predictable words were produced with longer durations than less predictable words for Korean provides evidence against a redundancy account repetition reduction, as well as prominence in general. In order to test these theories, it is important to try to replicate results across a number of different languages that have different linguistic properties. Effects that are linked to general production principles should be universal.

General Discussion

In the three experiments discussed in this chapter, both repetition and predictability influenced the prominence of target words in the tasks above. In English, repetition primarily affected word length and predictability primarily affected word intensity. In Korean, repetition affected both word length and F0, whereas predictability affected only duration. These results are most consistent with a multiple source view of acoustic prominence: the prominence of a word is affected by production factors like the lexical access account and by marking unpredictable information.

In fact, these findings are consistent with previous findings in the literature. Watson, Arnold, & Tanenhaus (2008) found a dissociation between intensity and duration in games of Tic Tac Toe, depending on the likelihood and importance of a game move. Baker & Bradlow (2009) have found that in clear speech, second mentions are more reduced for high frequency words than low frequency words, which is consistent with repetition and frequency effects having different underlying sources.

The duration results cannot be explained by an information theory account alone. These theories predict that speakers alter the duration of words such that listeners can more readily parse their utterances depending on the redundancy of the elements in the utterance. In particular, this account suggests that speakers should increase the length of less predictable words and reduce the length of predictable words. However, in Experiment 1a, when repeated words were unexpected, they were still reduced, despite being unexpected. In Experiment 2, where predictability and repetition were independently manipulated, predictability had no effect on target word duration. While predictability did have an effect in one of the measures of duration (target proportion), it was weaker than effects of repetition. Experiment 3 provides an even stronger argument against an information theory account of repetition. In Experiment 3, which was conducted in Korean, predictable words were actually produced with *longer* duration than less predictable words. This effect is in the opposite direction as predicted by information theory accounts of prominence reduction. Thus, it is clear that predictability alone is not sufficient to account for prominence differences in duration.

One potential concern is that the predictability manipulation here differs from the types of linguistic predictability that have typically been discussed in the literature. Linguistic predictability has been claimed to incorporate lexical and syntactic frequency, n-gram probabilities, and previous mention, all of which are properties of a language that the native speaker must learn through a lifetime of experience with the language. In contrast, the manipulation of predictability in the current study is based on the predictability of events occurring in a task. The representations that underlie this type of task-based predictability may differ from representations that underlie predictability based on stored linguistic experiences. However, stored sources of predictability are necessarily based upon input the speaker received

while interacting with his or her environment and language community. Thus, in principle, there is no reason why the manipulations of predictability in this task should differ from longer-term linguistic predictability, except that it is more recent.

Although the manipulation of predictability did not have a significant effect on the raw duration of the target word in English, this does not mean that speakers do not optimize some aspects of speech for processing by the listener. First, speakers did lengthen target proportions, a prediction made by information theoretical accounts. This can be done both by increasing the duration of the target word, or reducing the duration of words around the target word, with the end result being that the target word is perceived as being more prominent. Although this effect was relatively weak compared to effects of repetition, it may have resulted from some optimization of the signal for the listener. Second, speakers could still be providing a signal to listeners. In English, speaker may be improving word intelligibility by increasing the intensity of less predictable words. In Korean, even though speakers shortened less predictable words, this shortening of word duration could act as a cue to their listeners that a word is unexpected because the word is shorter than would be expected. Of course, the data here only speak to one kind of predictability, a very explicit cue based predictability. These data do not rule out the possibility that changes in duration can be explained by information theoretic accounts of other aspects of linguistic structure.

One question this chapter did not address is the locus of where these effects arise. Previous work suggests that the repetition effect is not realized at the phonological level. Repetitions lead to reduction, but saying a word and then its homonym does not (Fowler 1988). Homonyms are words that are identical in spelling and sound, but have different meanings. Because homonyms are identical in sound, the production process for homonyms should be

identical at the level of phonological encoding. This suggests that the repetition effect is situated at a higher level of production than phonological encoding, potentially at the level of lexical selection or message formulation. Fowler (1988) also failed to find a repetition effect when speakers produced repeated words that were produced as a list. Although the discrepancy between Fowler's (1988) findings and the results here are puzzling, one possible explanation is that word list production does not engage the same production processes, such as message planning, as situated language use. If reduction is linked to ease of processing at higher stages of the production process, one might not expect to see reduction in the production of word lists. At the very least, both of Fowler's findings suggest that effects of repetition may be driven by production factors earlier in the production process than the level of phonological encoding. In addition, work showing that reduction and decreases in intelligibility can occur even when a word is produced by a different speaker suggests that these effects are not necessarily rooted in phonological encoding (Bard et al., 2000; Anderson & Howarth, 2002). These are questions I will address in Chapter 5 of this dissertation.

It is also unclear at what level of production effects of predictability arise. One possibility is that there are feed-forward connections from the message formulation level to the level of articulation that modulate levels of intensity depending on the predictability of the word. Judgments of predictability could either come from explicitly modeling the expectations of the listener or by evaluating listeners' knowledge based upon the speaker's own assessment of predictability (e.g. Brown & Dell, 1987; Horton & Keysar, 1996). This is related to a broader debate in the psycholinguistics literature regarding the extent to which speakers design their utterances for the listener. Clearly, future work in this domain will need to determine the exact mechanism that underlies marking unpredictable information with intensity. More generally,

these data suggest that prominence is not a unitary linguistic or psychological construct. Different factors can play a role in whether a word is produced with prominence, and this prominence can be realized in different ways. In contrast to previous work, which has typically found increases in duration and intensity co-occurring in natural speech, I have found that intensity is more strongly linked to speaker expectation while duration is more strongly linked to repetition. These data suggest that the prominence of a word can potentially have multiple sources (Watson, 2010). The rest of this dissertation will be focused on exploring these different sources of prominence.

Chapter 3: The effect of addressee predictability on prominence

A central question about the language production system is whether speakers take into account the perspective of the addressee when designing utterances. This question applies to prominence as well. For example, while is it clear that less predictable words are produced with greater prominence than predictable words (Aylett & Turk, 2004; Bard et al., 2000; Bell et al., 2009; Fowler & Housum, 1987), there is some debate as to whether this effect is due to addressee design or due to speaker-centered processing.

The addressee design account argues that words that are easy for addressees to identify in context can be produced with reduced prominence whereas words that are less supported by the context are produced with greater prominence. According to this account, speakers reduce prominence for predictable words because they are easy for their addressees to identify (Fowler & Housum, 1987; Fowler, 1988). Support for this account comes from a series of experiments by Fowler (1988). Fowler (1988) found that repetition reduction only occurs when the same word was produced in a discourse context with meaningful prose. Repetition elicited from reading word lists did not lead to reduction. Additionally, repetition reduction did not occur when words were preceded by a homophone. Based on these results, Fowler (1988) argued that reduction reflects the speaker's estimate of the addressee expectations of a word.

The speaker-centered account argues that speakers reduce prominence for predictable information because the information is predictable for the speaker himself (Bell et al., 2009; Jurafsky et al., 2001). This account argues that prominence can be tied directly to cognitive difficulty such that words that require more cognitive effort will be produced with greater prominence. A recent study by Galati & Brennan (2010) seems to provide support for the processing account. In this study, speakers were asked to watch a cartoon video and then had to

narrate the cartoon twice to an addressee for a later memory test. Critically, speakers were either asked to narrate the story two times to the same listener or once to one listener, and another time to a different listener. They found that during the second telling of the story, speakers reduced the duration of key words regardless of whether the listener was the same listener or a different listener. That is, speakers did not adjust articulation of duration for listener needs. Rather, speakers seemed to have reduced duration simply whenever they were repeating key words.

Experiment 4

In this study, I examine whether or not speakers are sensitive to predictability for their addressee. In particular, I will compare prominence when speaking to friends to prominence when speaking to strangers. The experiment is a follow-up to Experiment 2 from Chapter 2 in which I manipulated predictability of the referent. The previous experiment did not include an addressee. Previous research suggests that friends have an advantage compared to strangers in comprehending their friends' speech (Fussell & Krauss, 1989). Fussell & Krauss (1989) argue that this is because the speech of friends is more predictable than the speech of strangers. In a more recent study, Savitsky, Keysar, Epley, Carter, and Swanson (2011) found that interlocutors exhibit more egocentrism when interacting with friends than with strangers. They argue that this happens because speakers overestimate the effectiveness of utterances when speaking to friends than when speaking to strangers. If this is true, then participants may use reduced prominence when speaking to friends, with whom they shared more experiences, than when speaking with strangers whom they have just met in the context of the experiment.

Method

Participants

Twenty-nine pairs of people participated in this experiment in exchange for $8 per hour of participation. Data from five pairs was excluded due to audio equipment failure. All participants were native speakers of English with normal or corrected to normal vision and hearing. Additionally, none of the participants were colorblind. Participants were grouped as pairs. One type of pairs included friends who came to the experiment together. The other type of pair included strangers who had not met until they arrived at the experiment.

Materials

This experiment was adapted from Experiment 2 of Chapter 2. This task differed from Experiment 2 in two ways. First, the task was a partner task so speakers needed to describe images to an addressee. The addressee had a screen with the same images and clicked on the appropriate images on her own screen. As in Experiment 2, the listener saw a gray circle indicating which object was most likely to flash. Similarly, the speaker saw a circle around the same object as the listener. However, instead of seeing a gray circle, speakers were made aware of when a target was less predictable for the addressee by seeing either a red or a green circle. A green circle indicated that the listener's cue was valid and the flashing object was predictable for the listener. A red circle indicated that the cue was invalid and that the flashing object was not predictable for the listener. I included this manipulation so that the speaker knew whether or not the target was predictable for his/her listener. In the previous experiment, there was no matcher, so it was not critical that the speaker know in advance whether or not a target was predictable for the listener.

Procedure

When subjects arrived they were randomly assigned to either the speaker or addressee role. Before beginning the task, participants were shown a video of other subjects completing the task. The video served both to inform the participants of the task as well as to prime the desired structure "The noun1 is shrinking … the noun2 is flashing." After viewing the video, the addressee moved into another room. The two rooms were connected via audio chat. Before beginning the actual experiment, the experimenter conducted a sound check and adjusted the microphone level. During the sound check speakers were asked to count from one to five slowly. Then the experiment adjusted the microphone gain such that the peak intensity after transformation did not result in audio clipping. Following the sound check, participants began the communication task.

The speaker's task was identical to Experiment 2. The speaker saw the same array of images as the speaker. After the speaker described the shrinking event, the matcher clicked on that object. Then the matcher saw a gray circle cuing the most likely object for the flashing event. Then the matcher waited for the speaker's description before clicking on the flashing object. After clicking on both objects, the matcher notified the speaker that she was ready to move onto the next trial and then clicked on a button to begin the next trial.

Predictions

If prominence is strategically modulated according to predictability for one's addressee, there should be two main effects. Speakers should produce less prominent words in the predictable condition than in the less predictable condition and they should also produce words with reduced

prominence when speaking to friends than with strangers. This is because speakers' utterances are easier for their friends to understand than for strangers to understand (Fussell & Krauss, 1989). If speakers modulate prominence solely due to processing difficulty, then there may be a main effect of predictability such that predictable words are reduced, but there should be no effect of partner type.

Results

Results were analyzed using multilevel linear regression with random intercepts. Likelihood ratio tests determined that random slopes did not significantly improve model fit. Additionally, repetition was included only as a fixed effect to improve model fit. Predictable words were produced with lower intensity than less predictable words ($t=2.29$; $p<0.05$). Additionally, speakers produced words with greater intensity when paired with a stranger than when paired with a friend ($t=3.17$; $p<0.001$). There were no significant interactions with intensity (see Figure 3.1).

Duration showed a significant predictability by friendship interaction ($t=2.75$; $p<0.01$). Predictable words were shorter in duration than less predictable words when speaking to strangers. However, predictability did not affect duration when speaking to friends. Additionally, there was a marginally significant main effect of predictability ($t=1.99$; $p=0.06$). Predictable words were produced with shorter duration than less predictable words (see Figure 3.2). The overall pattern of results for duration, intensity, and F0 is presented in Table 3.1.

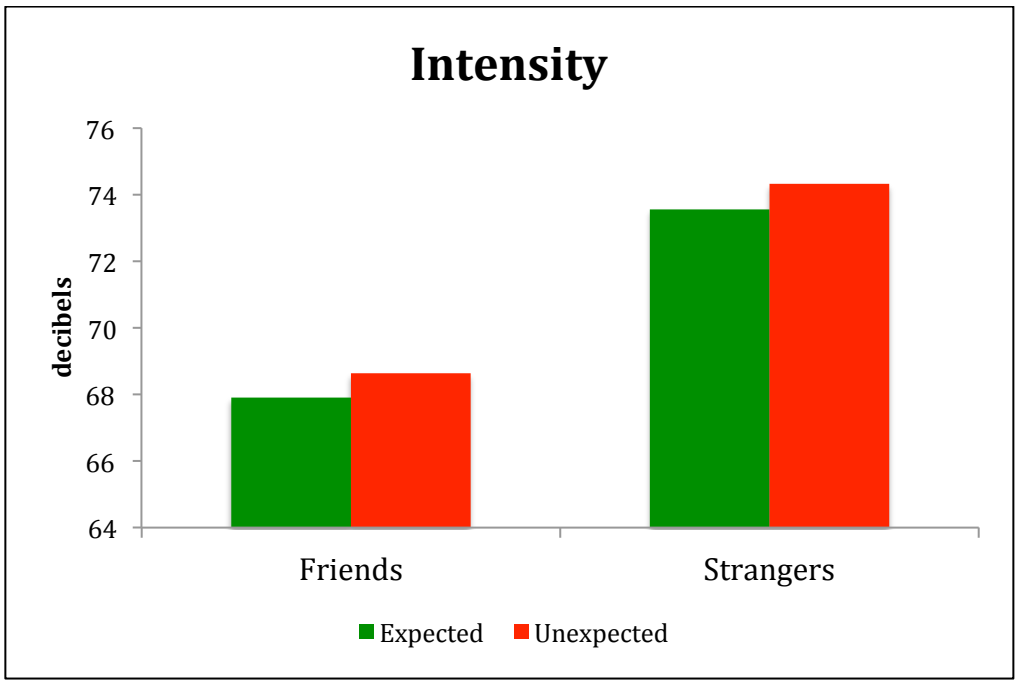

Figure 3.1. Presents the pattern for intensity across all four conditions of this experiment.

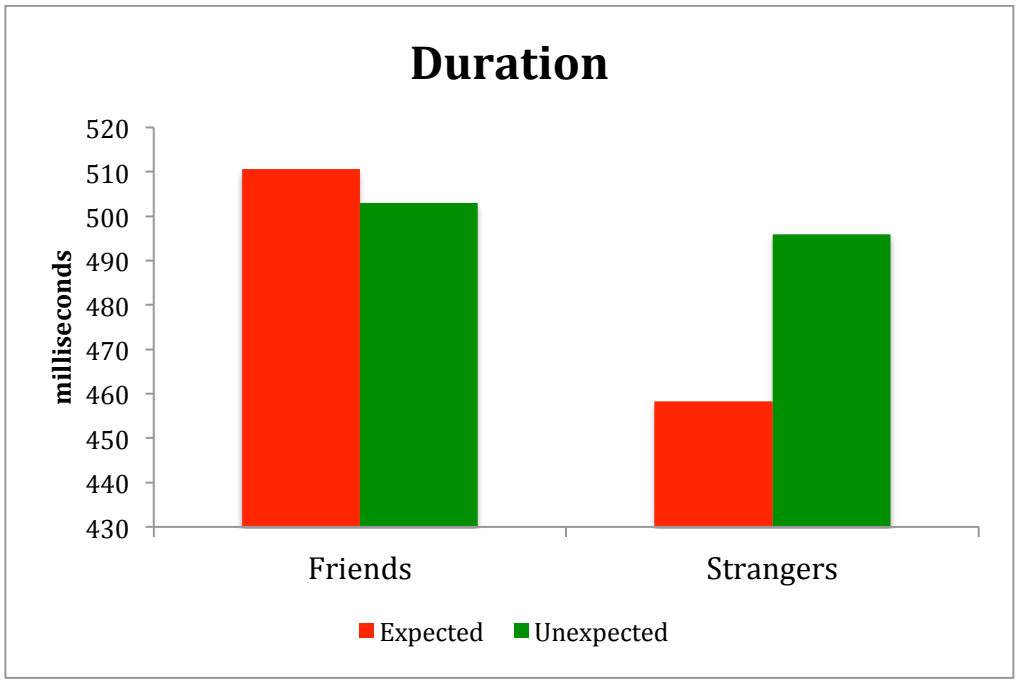

Figure 3.2. Presents the pattern for duration across all four conditions of this experiment.

Table 3.1.

	Friends		Strangers	
	Expected	Unexpected	Expected	Unexpected
Duration (ms)	511 (14.3)	503 (13.8)	459 (13.2)	496 (13.1)
Intensity (dB)	67.8 (0.482)	68.6 (0.538)	73.3 (0.539)	74.4 (0.555)
Average F0 (Hz)	186 (3.76)	187 (3.84)	167 (4.24)	173 (4.19)
F0 Maximum (Hz)	226 (6.92)	221 (5.96)	218 (8.32)	216 (7.71)
F0 Minimum (Hz)	162 (3.82)	167 (3.95)	141 (4.69)	149 (4.44)

Discussion

The pattern of results generally supports the idea that prominence is sensitive to one's addressee. Consistent with predictions of addressee design, speakers produced greater intensity when speaking to strangers than when speaking to friends and produced greater intensity in the less predictable condition than in the predictable condition. The pattern for duration was less consistent with addressee design theories. While speakers produced less predictable words with longer duration than predictable words, they did so only when speaking to strangers. However, the duration of utterances was longer when participants were speaking to friends than speaking to strangers. This pattern is inconsistent with audience design theories because words should be more predictable when speaking with friends and therefore they should be produced with less prominence. One possible explanation for this may be that there were inherent differences in the speech rate of the participants in the two conditions. Because this was a between subjects experiment, people in the friends condition are not the same people as in the strangers condition.

This is less of an issue for intensity because speaker baseline intensity was controlled for during the microphone sound check.

While this study showed that speakers produce different prominence patterns when speaking to friends than when speaking to strangers, there are a few methodological problems with this study. First, there was no control for duration. Duration was not measured independently of the partner manipulation, so it is difficult to know whether any differences in duration between friend and stranger pairs are due to the speakers themselves or due to the partner manipulation. Second, it is unclear if speakers reduced prominence because the presence of friends led to easier lexical access (e.g. Horton, 2007), or if speakers intentionally produced more prominent utterances when speaking to strangers (e.g. speaking style differences) perhaps because speakers felt they must speak more carefully when speaking to strangers (Savitsky, Keysar, Epley, Carter, and Swanson, 2011). If this reduction is due to ease of lexical access when speaking to friends, then the reduction is due to differences in the speaker's own production system. If speakers intentionally produce more prominent utterances when speaking to strangers (e.g. speaking style differences), then the reduction would be due to speakers explicitly designing their utterances for their addressees. One of the differences between these two accounts is that one is theory is based upon relatively fast priming of lexical access while the other theory assumes the act is somewhat intentional. Another difference is that the lexical access account relies on stored or prior common ground whereas the intentional design theory should be sensitive to local context that would lead speakers to assume more or less communicative efficiency. Stated differently, the intentional design theory would argue that the speaker controls how much emphasis he/she will place on words when design utterances for his/her addressee. Speakers may produce words with greater prominence if they are compelled to

design their utterances with greater prominence based upon whom they are speaking with. In Experiment 5, I will compare the effect of long-term friendship on spoken prominence to the effect of local familiarity.

Experiment 5

Following Experiment 4, which examined the effect of addressees in prominence production, I decided to run a more well controlled version of the experiment both to attempt to replicate the initial findings as well as to explain the pattern of results. Experiment 4 seemed to support the idea that speakers modulate their prominence with respect to their addressees' knowledge. Utterances from speakers in the friend condition were produced with lower intensity than utterances produced in the stranger condition. Additionally, friend pairs showed a weaker effect of predictability than stranger pairs. However friendship is a variable that encompasses a number of different possible factors that could lead to reduction. Experiment 5 attempts to test one of these factors: familiarity with one's partner. In Experiment 5, I will test whether or not familiarizing participants with their partners in a pre-experiment word generation task with no lexical overlap with experimental items can strengthen the partner effect. In addition, this experiment will control for baseline level prominence by having the participants read a passage from a novel before beginning the actual experiment. This baseline prominence will be used to control for differences in prominence that are inherent to the speakers themselves.

Method

Participants

The participants had the same characteristics as in Experiment 4. Participants were either pairs of friends who came at the experiment together or pairs of strangers who came to the experiment separately. All participants, both speakers and listeners, were compensated $16 in exchange for participating in this experiment. Thirty-two pairs of participants participated in this study.

Materials

This experiment had four parts: a passage reading task, a word generation task, a second passage reading task, and finally a referential communication task. The utterances from the first passage-reading tasks served as a baseline with which to compare speakers' average levels of prosody independently from the critical task manipulation. The passages were the first paragraphs from Chapters 2 and 3 of *War of the Worlds* (Wells, 1898). The referential communication task was identical to the one described in the Experiment 4 of this thesis.

There were two types of word generation tasks: a cooperative task and a non-cooperative task. Participants in the cooperative condition completed a modified version of the game Taboo. In this game, a person draws a card from a deck and must describe a target word or phrase that is printed on the card to his partner without saying the target word or phrase. Additionally, for each target word, there are taboo words that the speaker is also not allowed to mention. The partner uses these clues to guess the target word. Both the describer and the guesser can provide feedback to his partner and the guesser may continue guessing until s/he guesses the correct word. In this experiment, after the guesser correctly guessed the word, participants switched roles and the describer became the guesser. For the experiment, I removed all cards where the

target words and Taboo words were semantically related to the critical words in the referential communication task. I removed these cards in order to avoid lexical overlap between the target words in the cooperative pre-task. This game ended after six minutes. The goal of this pre-task was to familiarize the speaker with his/her partner. Participants in the non-cooperative task completed a category exemplar generation task. Participants were asked to write down as many exemplars as they could for each of three categories: flowers, types of dance, countries. Participants were given two minutes for each category before moving onto the next category for a total six minutes over the course of three categories. None of the critical items for the referential communication task were exemplars of these categories. Additionally, I collected their responses for the exemplar generation task and none of the responses overlapped with the critical items in the referential communication task.

Procedure

When subjects arrived they were randomly assigned to either the speaker or addressee role. After they had been assigned roles, the speaker completed a passage reading task while the addressee waited in another room. After reading the passage, both partners completed either a cooperative or a non-cooperative word generation task. Following the word generation task, the matcher left the room while the speaker completed another passage reading task. After the second passage reading task, the matcher joined the speaker in the experiment room and then both participants began the referential communication task.

Predictions

If addressee effects on prominence are due to familiarity with the partner, then speakers should produce words with less prominence when speaking to friends than when speaking to strangers. Additionally, if speakers accommodate to familiar partners, there may be an effect of pretask cooperativeness because speakers may reduce prominence when speaking to partners with whom they completed the cooperative task because they are more familiar with them, but not from partners who did the non-cooperative task. Finally, the pretask manipulation could interact with friendship such that only stranger pairs show a pretask cooperativeness benefit while friends are at ceiling for reduction.

Results

Of the 32 pairs of participants, four pairs were excluded from analysis. Three of these were friend pairs and one of these was a stranger pair. The first friend pair was excluded because of a programming error in the experiment, which led to the experiment not being able to be completed. One friend pair was excluded because the speaker described the cuing event as well as the flashing event leading the speaker to mention the target multiple times. This also caused the listener to be misled into believing that the cued object was a good cue. The third friend pair was excluded because the speaker used pronouns to describe objects in the repeated condition. One stranger pair was excluded because the speaker was not a native speaker of American English.

The data were analyzed using linear mixed effects regression with centered predictor variables. Likelihood ratio tests determined that random slopes did not significantly improve model fit for any of the models tested. I will report duration with three metrics, raw duration of

the target word, target to utterance duration proportion (i.e. target duration/utterance duration), and a ratio of raw value to mean value from the reading task (raw duration/ mean duration). I will report intensity and F0 with two different metrics: raw value of the target word, and a ratio of raw value to mean value from the reading task (raw prominence/ mean prominence). P-values were obtained using Markov Chain Monte Carlo (MCMC) sampling.

Raw duration showed two main effects: repeated words were shorter than non-repeated words ($t=10.94$, $p<0.0001$) and predictable words were shorter than less predictable words ($t=2.68$, $p<0.05$). Target proportion also showed a main effect of repetition: repeated words were produced with a lower noun to utterance duration proportion than non-repeated words ($t=7.61$, $p<0.0001$). Target proportion showed a marginally significant effect of predictability: predictable words were produced with a lower noun to utterance duration proportion than less predictable words ($t=1.73$, $p<0.09$). Finally, target proportion showed a significant effect of pretask: speakers in the cooperative condition produced sentences with a smaller target proportion than speakers in the non-cooperative condition ($t=2.16$, $p<0.05$). Raw duration to mean duration ratio showed two main effects: repeated words were produced with a smaller raw duration to mean duration ratio than non-repeated words ($t=10.35$, $p<0.0001$) and predictable words were produced with a smaller raw to mean duration ratio than less predictable words ($t=2.51$, $p<0.05$). Additionally, raw to mean duration ratio showed a main effect of friendship status: friends produced utterances with a smaller raw to mean duration ratios than strangers ($p=2.68$, $p<0.05$). Figures 3.3 – 3.8 show the means for all the duration measures.

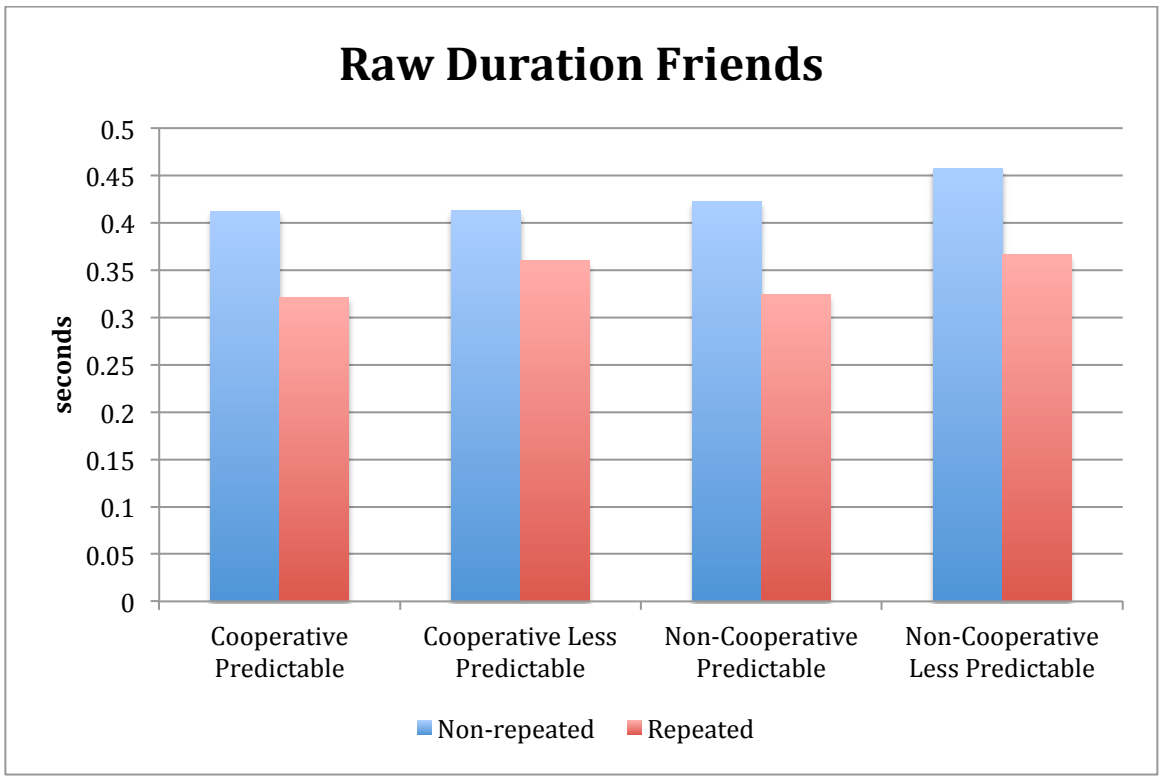

Figure 3.3. Presents the averages for raw duration for speakers who are partnered with a friend.

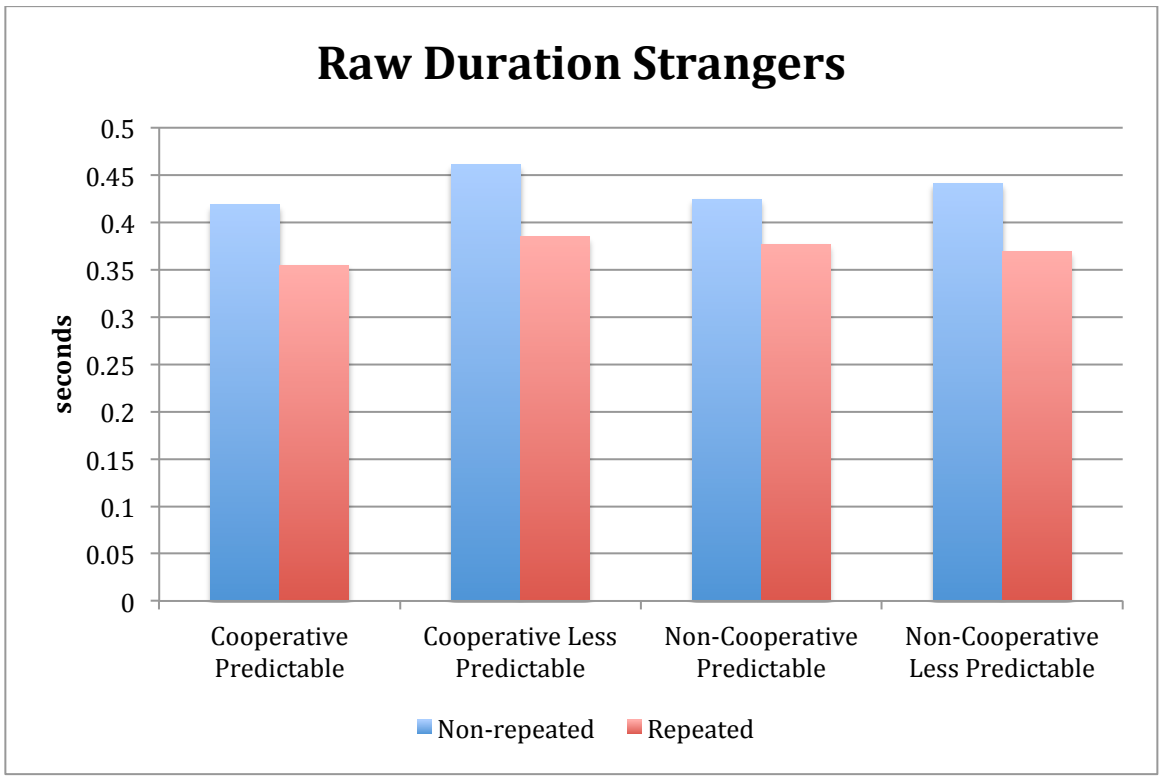

Figure 3.4. Presents the averages for raw duration for speakers who are partnered with a stranger.

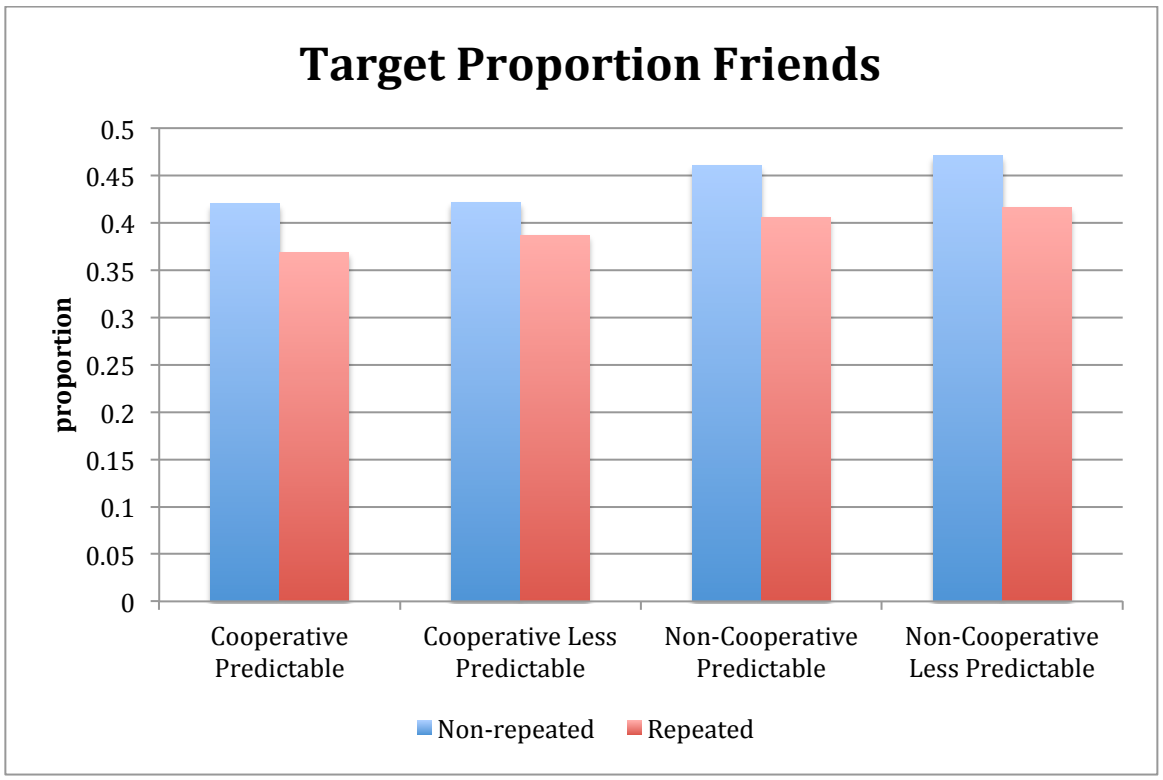

Figure 3.5. Presents average proportion of the utterance that is covered by the target word for speakers who are partnered with a friend.

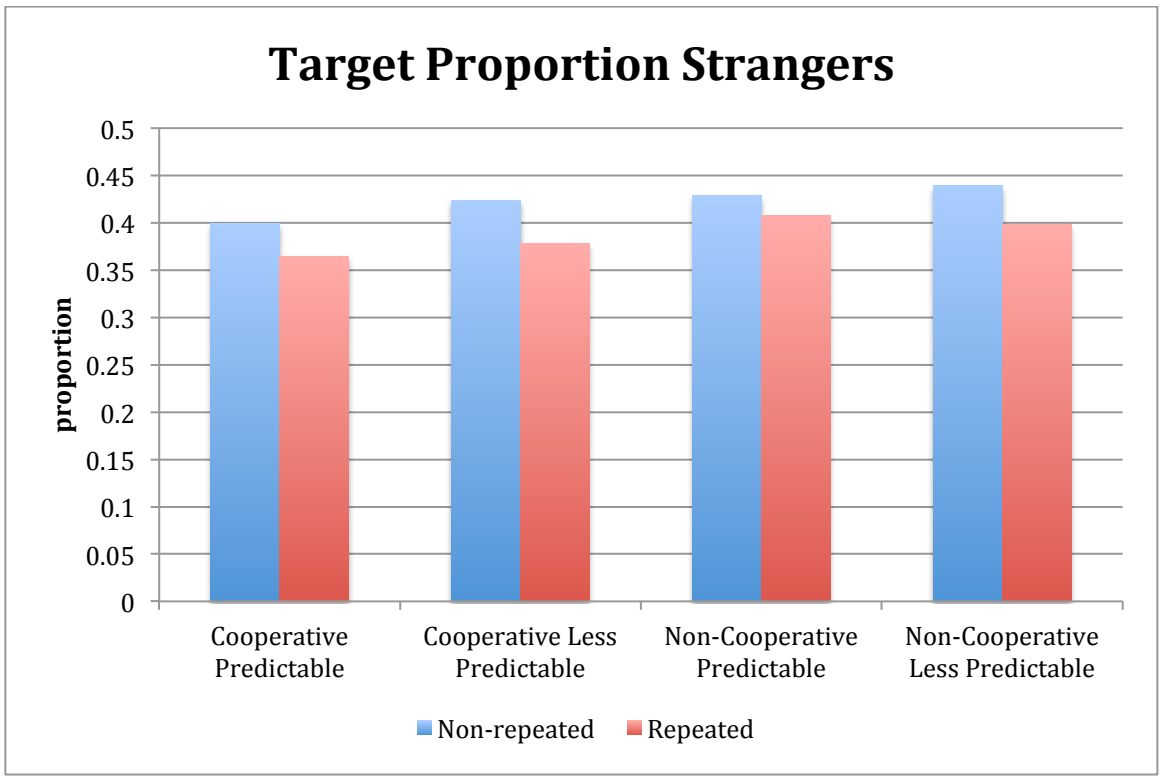

Figure 3.6. Presents the average proportion of the utterance that is covered by the target word for speakers who are partnered with a friend.

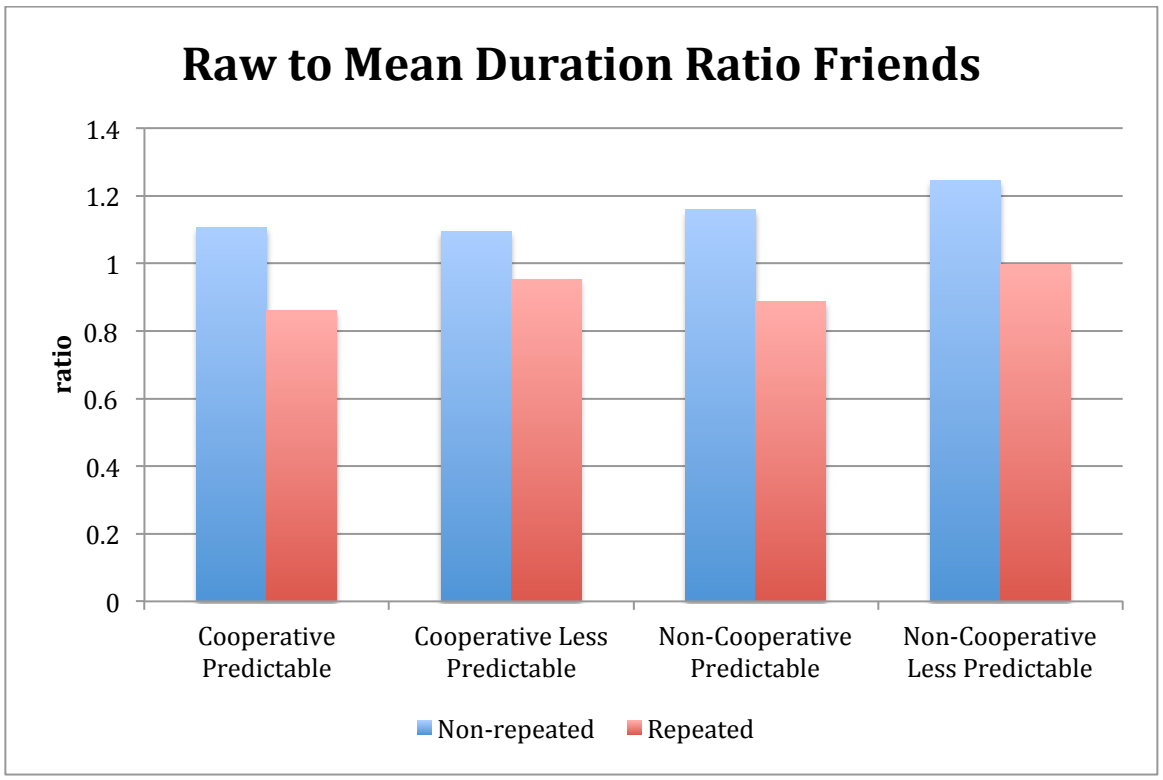

Figure 3.7. Presents the average ratio of target duration to mean duration for speakers who are partnered with a friend.

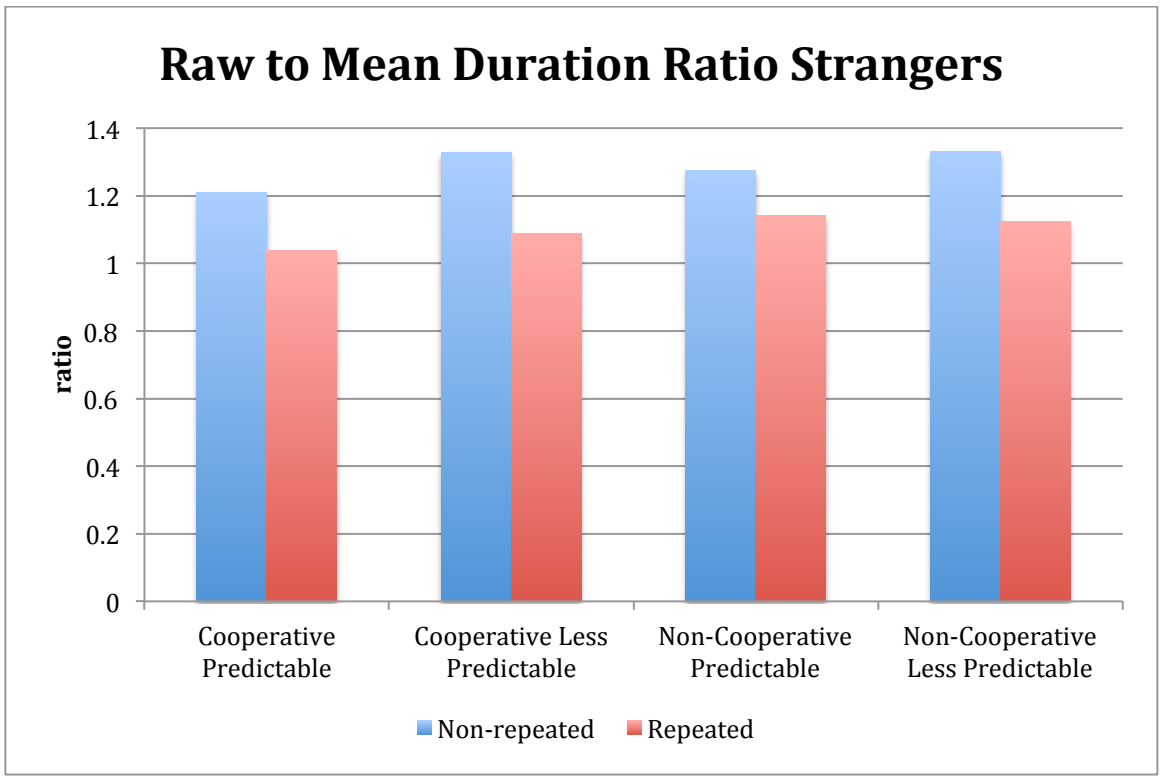

Figure 3.8. Presents the average ratio of target duration to mean duration for speakers who are partnered with a stranger.

Like raw duration, raw intensity showed two main effects: repeated words were produced with lower intensity than non-repeated words ($t=2.94$, $p<0.01$) and predictable words were produced with lower intensity than less predictable words ($t=2.89$, $p<0.01$). There was also a marginally significant interaction of pretask and repetition. Speakers in the non-cooperative condition had a smaller effect of repetition than speakers in the cooperative condition, ($t=-1.97$, $p=0.06$). There was also a marginally significant interaction between repetition and predictability. Words in the predictable condition showed a stronger effect of repetition than words in the less predictable condition ($t=-1.87$, $p=0.08$). Finally there was a marginally significant 3-way interaction between friendship, repetition, and predictability. The previously mentioned

repetition by predictability interaction was stronger in stranger pairs than in friend pairs ($t=1.95$, $p=0.07$). Raw intensity to mean intensity ratio showed significant main effects of repetition and predictability: repeated words were produced with a lower raw to mean intensity ratio than non-repeated words ($t=2.83$, $p<0.01$) and predictable words were produced with a lower raw to mean intensity ratio than less predictable words ($t=2.86$, $p<0.01$). There was also a significant pretask by repetition interaction. Speakers in the cooperative condition show a stronger effect of repetition than speakers in the non-cooperative condition ($t=-2.01$, $p<0.05$). Finally there was a significant 3-way interaction between friendship status, repetition, and predictability. Friends showed main effects of repetition and predictability in the same direction as the simple main effects while strangers showed a repetition by predictability interaction such that in the non-repeated condition predictable words were produced with a lower raw to mean intensity ratio than less predictable words while in the repeated condition, predictable and less predictable words were produced with similar raw to mean intensity ratios ($t=2.33$, $p<0.05$). Figures 3.9 – 3.12 show the means for the intensity measures.

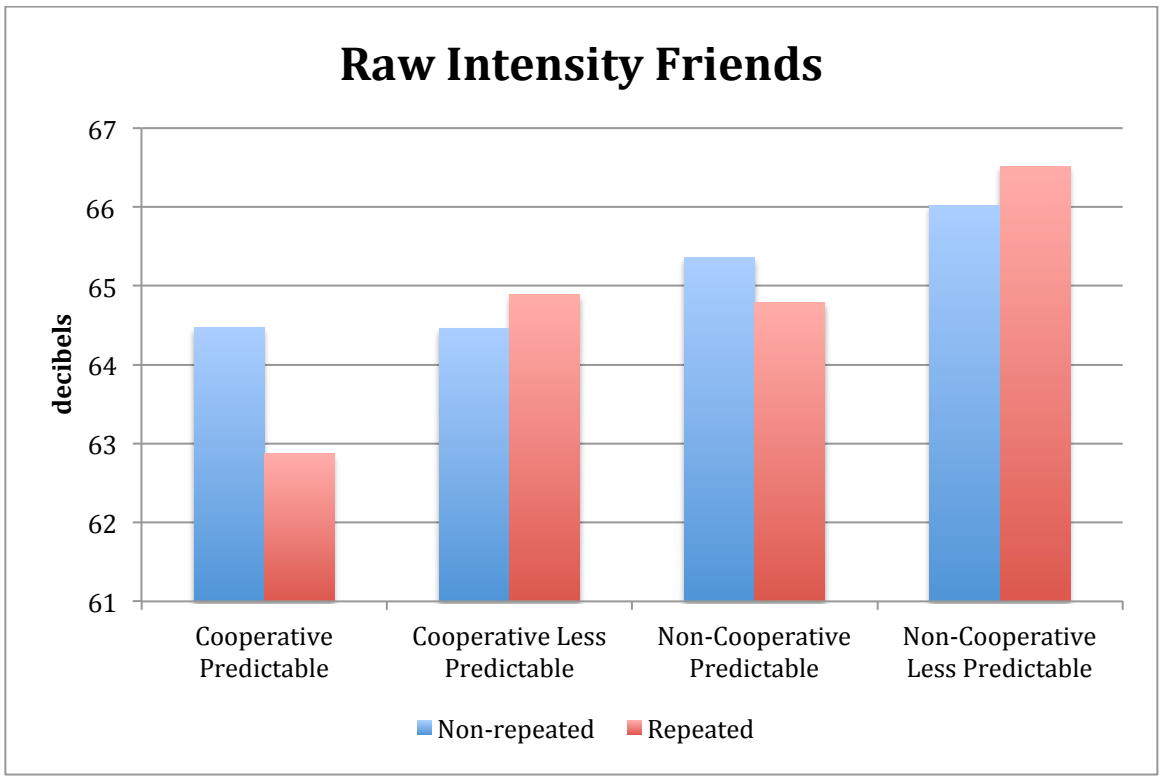

Figure 3.9: Presents the raw intensity for speakers who are partnered with a friend.

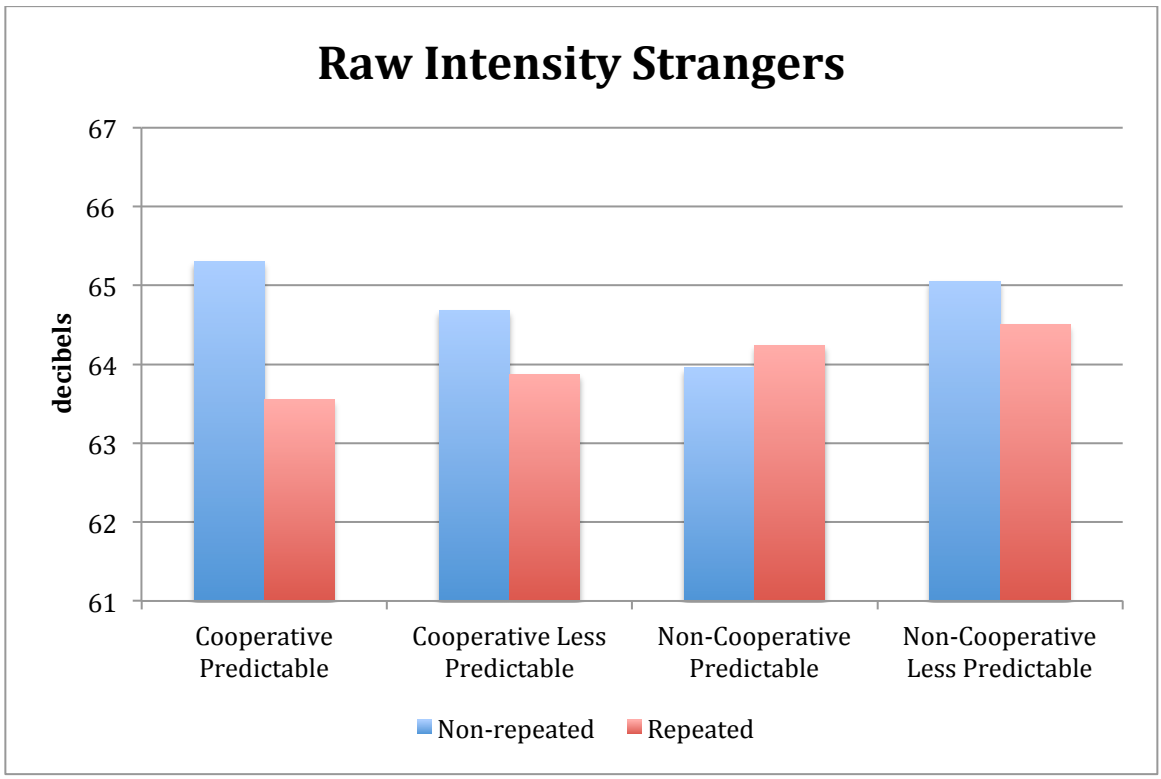

Figure 3.10: Presents the raw intensity for speakers who are partnered with a stranger.

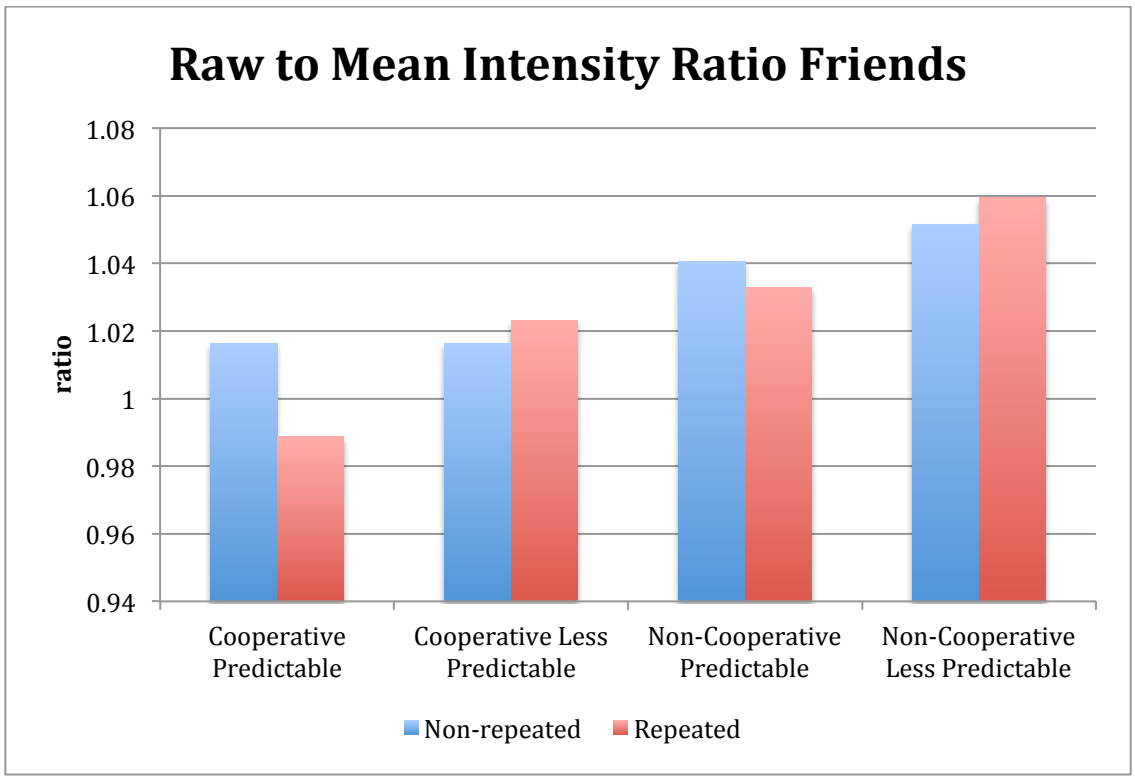

Figure 3.11. Presents the average ratio of target intensity to mean intensity for speakers who are partnered with a friend.

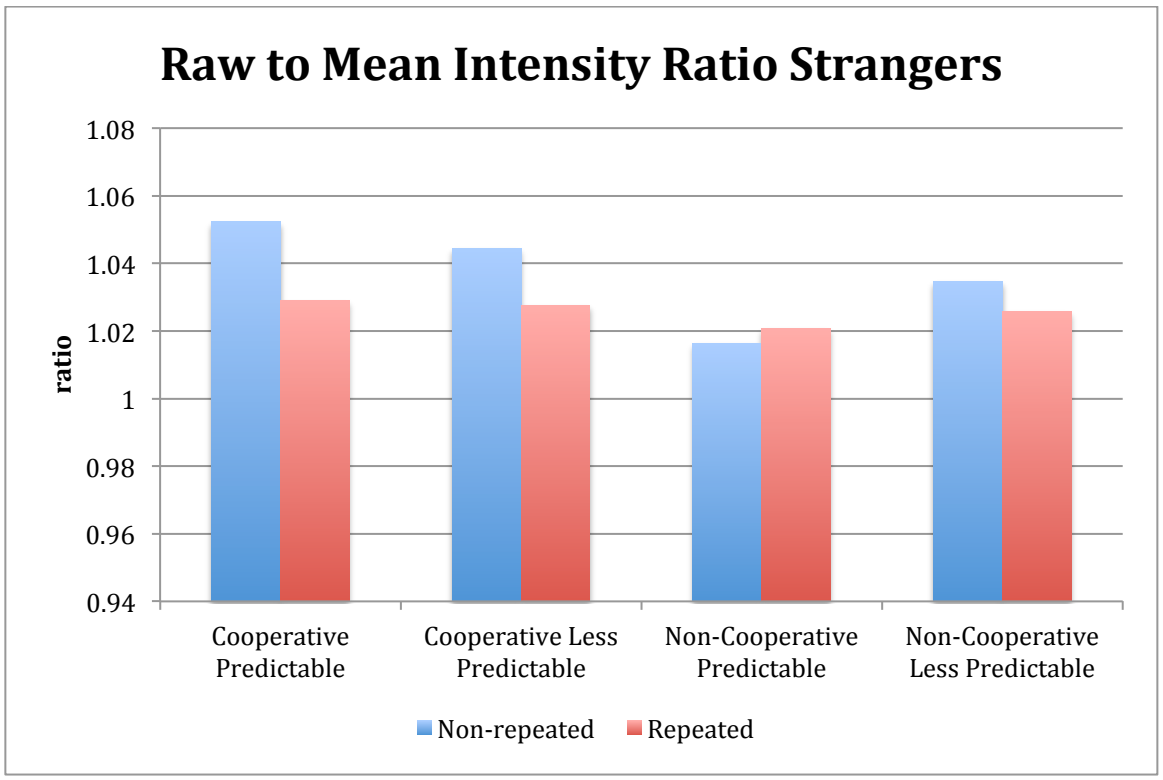

Figure 3.12. Presents the average ratio of target intensity to mean intensity for speakers who are partnered with a friend.

Raw F0 showed only one main effect of pretask. Speakers in the cooperative condition produced words with higher F0 than speakers in the non-cooperative condition ($t=-2.68$, $p<0.01$). There was also a significant interaction between friendship and pretask. Friend pairs had a larger effect of pretask than stranger pairs ($t=1.58$, $p<0.01$). Finally, there was significant 3-way interaction between pretask, repetition, and predictability ($t=-2.46$, $p<0.05$). In the cooperative condition, repeated words in the predictable condition were produced with higher F0 than repeated words in the less predictable whereas non-repeated words in the predictable condition were produced with lower F0 than non-repeated words in the less predictable condition. In the non-cooperative condition, there were two main effects: repeated words were produced with

lower F0 than non-repeated words and predictable words were produced with lower F0 than less predictable words. Raw to mean F0 ratio showed a marginally significant interaction of friendship status and predictability. For friend pairs, predictable words had a lower raw to mean pitch F0 than less predictable words while for strangers, there was no difference in raw to mean F0 ratio for predictable and less predictable words ($t=-1.90$, $p=0.06$). There was also a significant 3-way interaction between pretask, repetition, and predictability ($t=-2.07$, $p<0.05$). The pattern was the same as the 3-way interaction on raw F0 values. Figures 3.13 – 3.16 show the means for the F0 measures.

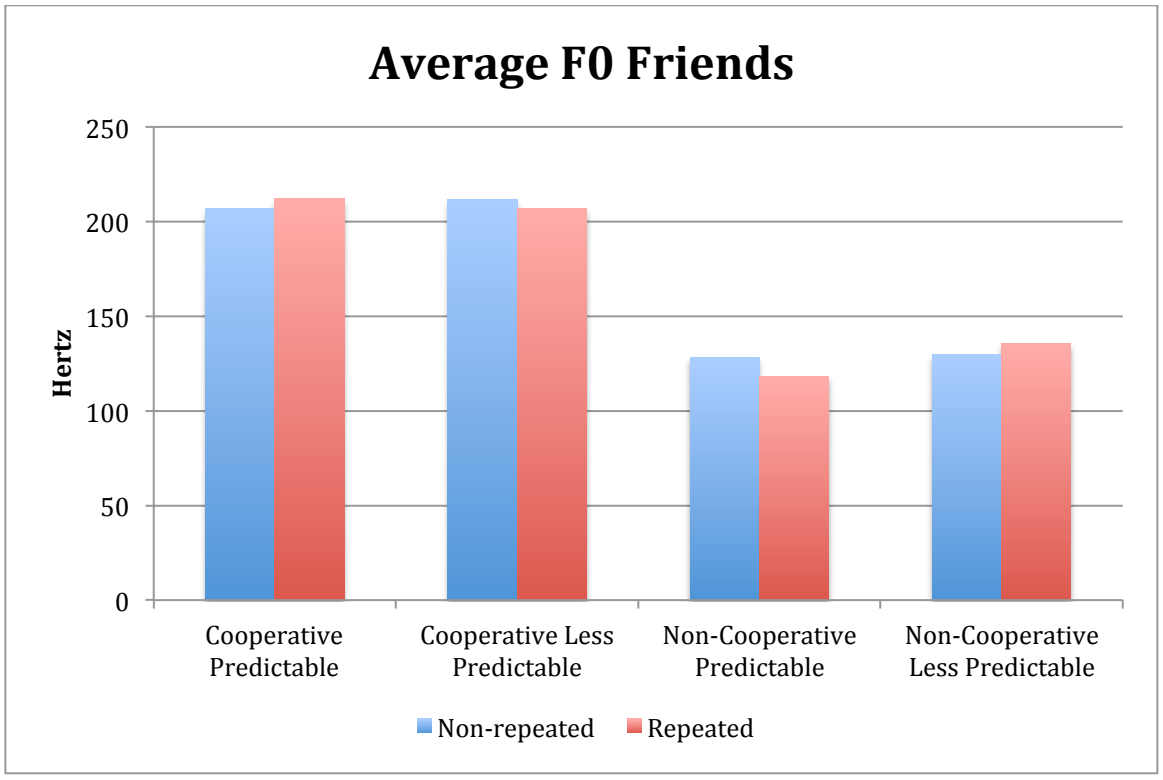

Figure 3.13. Presents the raw average F0 for speakers who are partnered with a friend.

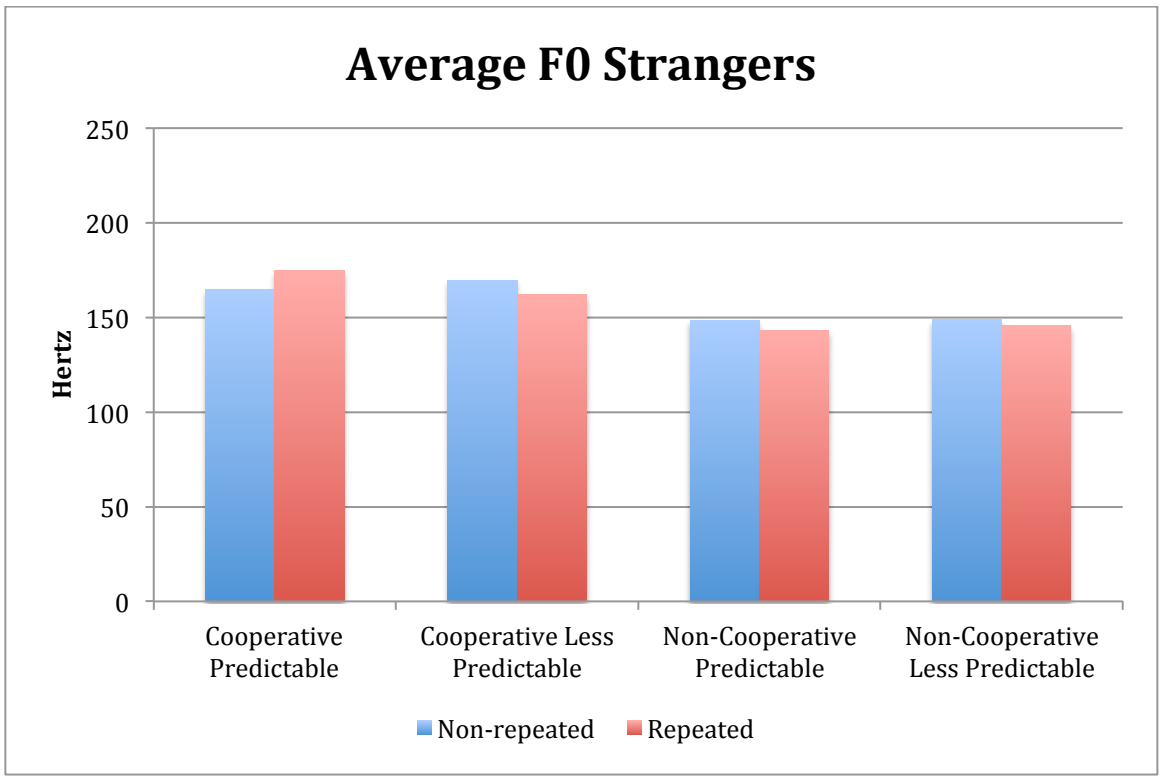

Figure 3.14. Presents the raw average F0 for speakers who are partnered with a stranger.

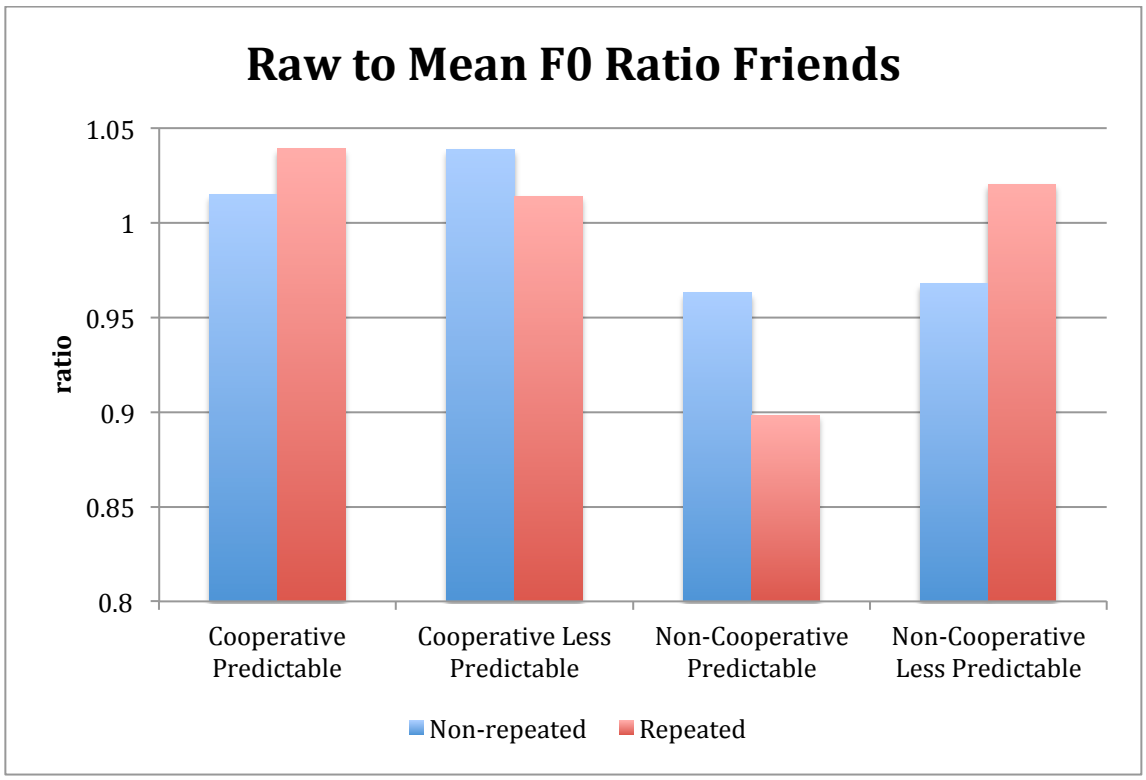

Figure 3.15. Presents the average ratio of target F0 to mean F0 for speakers who are partnered with a friend.

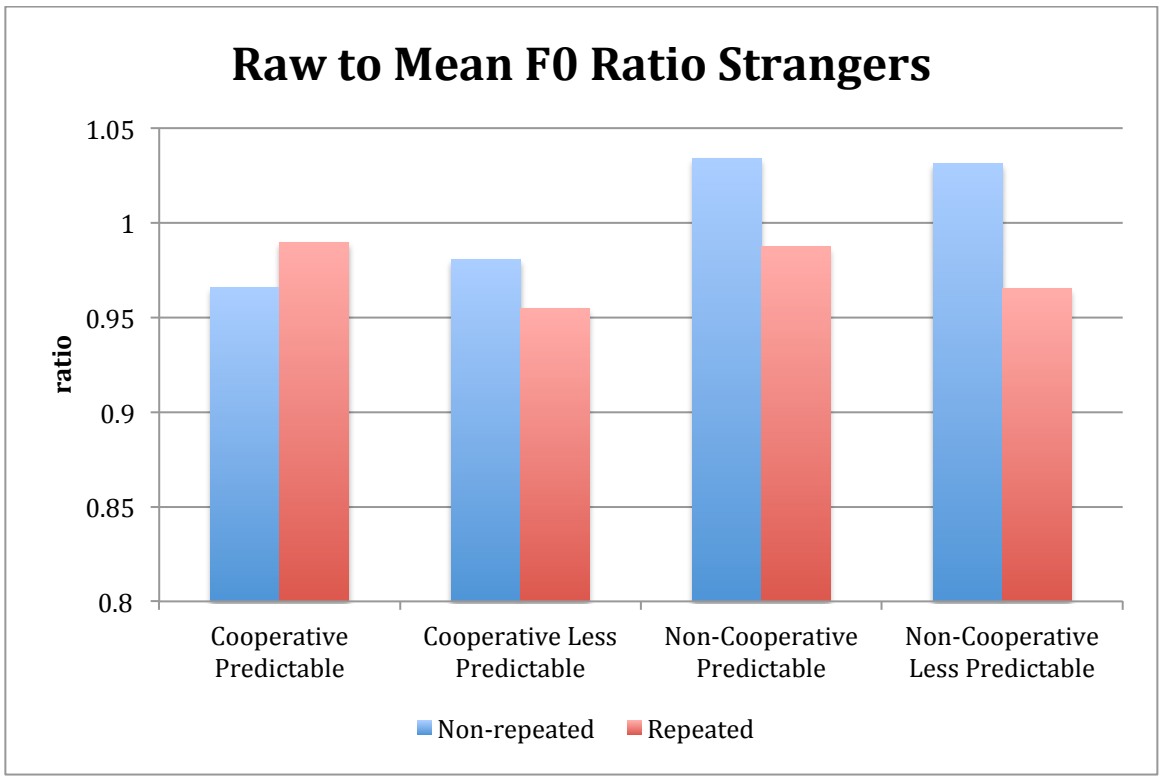

Figure 3.16. Presents the average ratio of target F0 to mean F0 for speakers who are partnered with a stranger.

Discussion

In general, the results from Experiment 5 did not replicate the patterns from Experiment 4 with respect to addressee type. Unlike Experiment 4, there was no significant main effect of friendship on intensity, although there was a main effect of friendship on raw to mean duration ratio. This suggests that the differences in Experiment 4 may have been due to random noise the in condition assignment. The speakers in the stranger pair condition in Experiment 4 may simply have had louder voices than the speakers in the friend pair condition. Also unlike Experiment 4, there was no significant interaction between friendship and predictability. When friendship did interact with prominence, it usually affected both repetition and predictability in a 3-way interaction. I will discuss this in more detail in the general discussion of this chapter.

The pretask manipulation showed clearer patterns than the friendship manipulation. Speakers in the cooperative pretask had shorter noun to utterance ratios than speakers in the non-cooperative condition. In this sense, speakers in the cooperative condition produced less prominent utterances than the speakers in the non-cooperative condition. However, for all other variables, the pretask manipulation seemed to have led to greater prominence in the cooperative condition. Speakers in the cooperative condition had a greater effect of repetition on intensity than speakers in the non-cooperative condition. Finally speakers in the cooperative condition produced utterances with greater F0 than speakers in the non-cooperative condition. One possibility is that the pretask manipulation has multiple underlying factors, an implicit one and an explicit one. Perhaps duration was reduced due to participants being familiar with speaking to their partner. This would be consistent with the ease of processing results from Chapter 2 that suggests that the repetition effect on duration is due to ease of processing. At the same time, speakers may have been more motivated to produce informative utterances after having completed a cooperative pretask than after completing a non-cooperative pretask because the cooperative pretask requires participants to work together. As a result, speakers in the cooperative condition produced words with greater intensity and F0. This interpretation is consistent with Watson, Arnold, & Tanenhaus (2008), which showed that informative words are produced with greater intensity than less informative words. Regardless of the direction of the effects, the results suggest that prominence can be affected by familiarizing a speaker with his/her interlocutor.

General Discussion

Across a pair of experiments I examined the effect of addressee on prominence production. In Experiment 4, there was a main effect of friendship such that speakers in friend pairs produced utterances with lower intensity than speakers in stranger pairs. Moreover, in Experiment 4, there was an interaction of friendship by predictability for noun duration. These effects were not replicated in Experiment 5. There could be a number of reasons why the friendship pattern did not replicated. One possibility is that it is due to the differences between Experiment 4 and Experiment 5. Experiment 5 had a reading task as well as a word generation pretask while Experiment 4 had neither of these tasks. It is possible that by conducting the reading task and the cooperative or non-cooperative pretask, speakers maybe have changed the way they interacted with their partners. Given the fact that the pretask seems to have affected speakers' prominence, it is possible that the pretask may have affected how friend pairs approached the task. One possible explanation for the effect in Experiment 4 is that friend pairs were speaking to each other right before the experiment because they arrived together for the experiment. This may be similar to the cooperative condition from Experiment 5. It is possible that the friendship effect from Experiment 4 is really due to a local effect of familiarity because friends had just been speaking with each other whereas stranger pairs had not. In Experiment 5, because of the pretask, it is possible that the non-cooperative condition weakened this familiarity while the cooperative condition maintained the effect.

While the friendship effects were not replicated across the two experiments, the fact that the pretask manipulation seems to have affected prominence suggests that prominence is sensitive to effects of familiarity with one's partner. The effect of pretask taken together with the

lack of a friendship effect in Experiment 5, suggests that local context may have a greater effect on prominence production than long-term stored relations such as friendship.

Chapter 4: Repetition reduction: dissociating form repetition and reference repetition

As discussed in Chapter 2, there is a wealth of evidence showing that repeated words are produced with reduced prominence. Though this effect is typically discussed in terms of referents being either new or given in a discourse, there is some ambiguity as to what it means to be repeated. Repetition can mean a number of different things. It can include repeatedly mentioning the same referent, as in the case of pronoun use. This would be repetition at the message level. Or repetition could refer to the repetition of a lexical item: repeatedly using the same word. This would be repetition at the level of lexical selection. Repetition reduction, therefore, could arise from multiple levels of the production process. The question of how repetition is represented in the production system is of particular interest because it may reveal details about how prominence is more generally represented in the production system.

In a previous study, Fowler (1988) argued that the repetition effect is likely rooted at the message or discourse level of production. In her study, she asked participants to read aloud paragraphs that contained critical words. Before the target sentence, participants read a sentence that contained either a homophone of the target word, a previous mention of the target word that shared its meaning, or an unrelated word. She found that while repetition of the same word led to shorter duration, words that were preceded by a homophone were not produced with shorter duration. This pattern suggests that simply repeating the articulatory pattern for a word is not enough to lead to reduced prominence. However, Fowler (1988) never tested whether it is repetition of the referent itself that matters, or simply repetition of a particular referring expression (i.e. lexical form repetition).

In this chapter, I test whether repetition reduction is due to repeatedly accessing a particular *referring expression*, or due to repeatedly mentioning a particular *referent*. A study by Kahn & Arnold (2010) provides some evidence that such a distinction is warranted. In their study, participants described an object undergoing a change. The object was either new, linguistically given (previously mentioned), or non-linguistically given (visually cued but not previously mentioned). They showed that both linguistically given and non-linguistically given words are produced with reduced prominence, however linguistic givenness leads to greater reduction than non-linguistic givenness.

It is also possible that these two types of repetition may affect prominence in different ways. This would be consistent with what Watson (2010) calls a multiple source theory of prominence. According to Watson (2010), prominence is best explained as the result of a combination of effects from multiple sources instead of a single source. According to the multiple source view of prominence, factors that are typically correlated in natural speech may have independent effects on prominence and may affect the acoustic signal in different ways. For example, Experiment 2 from Chapter 2 showed that while repetition and predictability are correlated in natural speech, predictability is more strongly linked to changes in intensity whereas repetition is more strongly linked to changes in duration. Thus, under a multiple source account of prominence, referent repetition and lexical repetition may both affect prominence, but affect different aspects of the acoustic signal.

Experiment 6

In Experiment 6, I explored the repetition effect by testing whether repetition reduction is due to repeated mention of a referent independent of the form of the referring expression, or due

to repetition of a referential form independent of the intended referent. Speakers described two events involving different characters with different occupations moving between locations on a computer screen. Sometimes the same referent was mentioned twice. Other times, one referent was mentioned and then a different referent was mentioned. Additionally, sometimes the two events involved characters that were either described with the same referring expression or with different referring expressions.

Method

Participants

Sixteen people participated in this study. Participants were a mix of subject pool participants from the University of Illinois and paid subjects living in Champaign or Urbana, Illinois. Paid participants were compensated $8 for one hour of participation. Participants recruited from the subject pool were compensated with course credit in exchange for participation. All participants were native speakers of American English with normal or corrected to normal vision and hearing.

Materials

In order to create a context in which the same referent could be referred to using different referring expressions, I created an event description task in which participants described images of characters who were training for different occupations at two different training centers. The occupations of the characters on the screen were all identifiable from their clothing. The training centers were at the University of Illinois and Parkland College. Both of these institutions are located in the Champaign-Urbana area.

The images used in this experiment were created using the Sims creator from © The Sims 3, a video game. Because of the level of detail in both faces and outfits, subjects were able to identify individuals as well as their occupations. I created 14 different characters with varying physical characteristics. Eight of the characters were male, and six characters were female.

There were two factors: repetition of referents and repetition of referring expressions (form). This yielded four conditions: Non-repeated character, non-repeated occupation; Non-repeated character, repeated occupation; Repeated character, non-repeated occupation; Repeated character, repeated occupation (See Figure 4.1 for an example display of each condition). In order to have enough statistical power, items were repeated over the course of the experiment such that each occupation was used two times as a critical word, but with different characters across two blocks of trials. As a result, I also tracked the block in which the participant encountered a particular occupation as a critical target. The block manipulation was hidden from the participants.

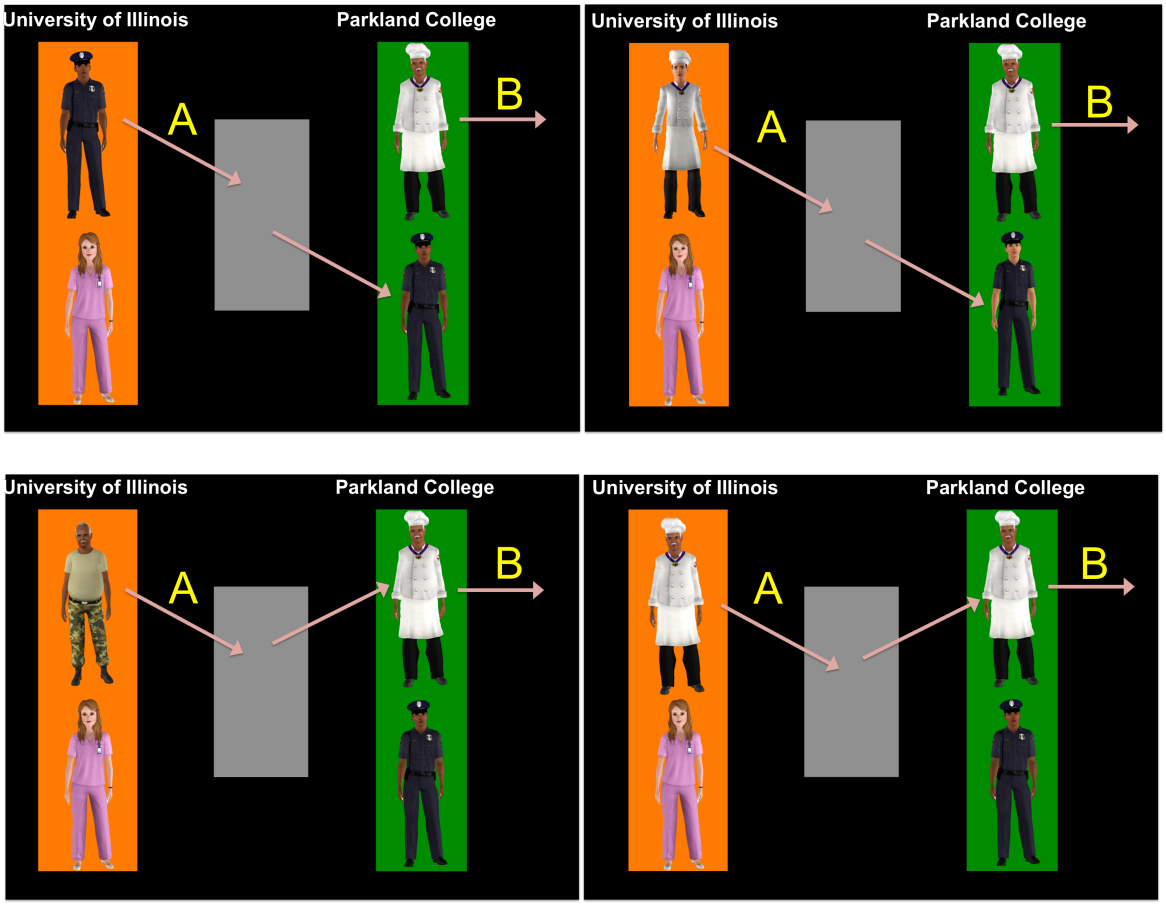

Figure 4.1. Depiction of the four conditions from this experiment. The top left image shows the different referent, different form condition. The top right image shows the different referent, same form condition. The bottom left image shows the same referent, different form condition. The bottom right image shows the same referent, same form condition. The arrows represent the events on the screen. Speakers described the events at point A and point B.

Images were displayed using Matlab with the Psychophysics Toolbox version 3. At the beginning of a trial, participants saw an area labeled University of Illinois on the left side of the screen and a grey, "home" area on the right side of the screen. At the beginning of each trial, there were two characters with different uniforms at the University of Illinois. Then, one of these

characters moved to the home region. The participant described this leaving event. Then the screen panned left such that the University of Illinois was no longer visible but the "home" and Parkland College were visible. The Parkland College training center contained a third character. On half of the trials, the character at "home" moved over to Parkland College while wearing the same outfit s/he was wearing at the University of Illinois. Then one of the two characters at the Parkland College training center moved off screen. The participant described the second leaving event. On the other half of the trials, the character at "home" briefly disappeared and reappeared on screen in a different uniform before moving to Parkland College. Then, one of the two characters at Parkland College moved off screen. The sequence of events is depicted in Figure 4.2. Example 4.1 shows the types of sentences produced in each condition.

Example 4.1.

Different referent, different name: "The doctor (man1) is leaving A. The detective (man2) is leaving B."

Different referent, repeated name: "The detective (man1) is leaving A. The detective (man2) is leaving B."

Repeated referent, different name: "The doctor (man1) is leaving A. The detective (man1) is leaving B."

Repeated referent, repeated name: "The detective (man1) is leaving A. The detective (man1) is leaving B."

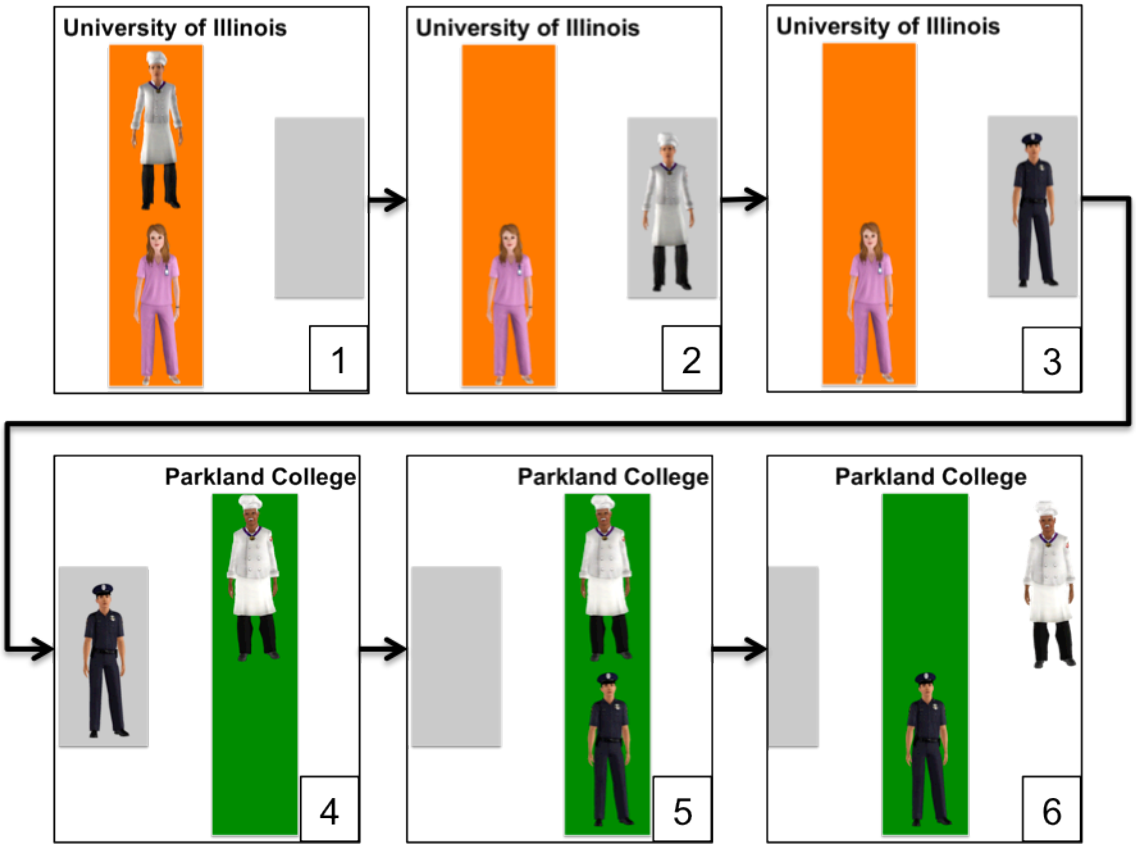

Figure 4.2. Depicts the sequence of events for the different referent, same form condition as they occur on the participant's screen.

Because some of the occupations were more readily identifiable through uniforms than others, participants were trained on the names of the sixteen occupations before beginning the experiment. During training, participants were shown a picture of a character wearing his or her occupation's uniform. The occupation name was displayed below the picture. Participants were required to say the occupation name aloud. After naming all of the occupations once, they were presented with test trials in which the image was display without the occupation name. Participants again had to name the images. If the participant had forgotten the name, the experimenter gave the participant the name.

Following the name training, participants were presented with four practice trials, one from each condition. After completing the practice trials, participants completed the 32 critical trials. There were no filler trials. Participant descriptions were recorded using a microphone headset. The microphone was positioned at a constant distance from the mouth of the speaker in order to accurately measure the intensity of the sound wave. Praat was used to extract intensity and duration values from the target word, which was the occupation name in the second utterance on each trial.

Predictions

If repetition reduction arises from repetition of the referent, independent of the referring expression used, then repeated referents should be reduced even when the occupation name has changed. If repetition reduction is due to repetition of the referring expression (lexical form), then repeated occupation names should lead to reduction, even when the referent has changed. Additionally, it may be possible that both kinds of repetition lead to reduction independently. This would be consistent with a multiple source account of prominence (Watson, 2010).

Results

The data were analyzed using multilevel mixed effects regression with random slopes and intercepts. Using likelihood ratio test, it was determined that the best models included only random intercepts with no slope term. The results are discussed in terms of duration and intensity. For duration, there was a strong main effect of form repetition such that repeated forms had shorter durations than non-repeated forms ($t=6.716$; $p<0.001$); however, there was no effect of referent repetition. There was also a significant main effect of encounter such that occupation

names were produced with shorter duration on the second encounter ($t=-3.994$, $p<0.001$). Additionally, there was a three-way interaction between form, referent, and number of encounters ($t=2.130$; $p<0.05$). On the first encounter, words from the same referent, different form condition were produced with longer duration than all other conditions. The duration patterns are shown in Figure 4.3.

As with duration, there was a significant main effect of form repetition on intensity such that repeated forms led to lower intensity ($t=3.62$, $p<0.01$), but there was no main effect of referent repetition. Unlike duration, there was no main effect of number of encounters. There was also a three-way interaction between form, referent, and encounter for intensity, however the pattern of the interaction was different from the pattern for duration ($t=-2.41$, $p<0.05$). In the first block, in the different referent condition, if forms were repeated, then the word was reduced in intensity compared to when the forms were different. However, when the referent was the same, intensity did not differ across the same form and different form conditions. The pattern for the second block was different from the pattern in the first block. In the second block, words from the repeated referent, repeated form condition were produced with intensity that was even lower than words from the different referent, repeated form condition. This pattern suggests some special status for referents that are repeated with exact same referring expression such that referent repetition only matters when the form is also repeated. The intensity patterns are shown in Figure 4.4.

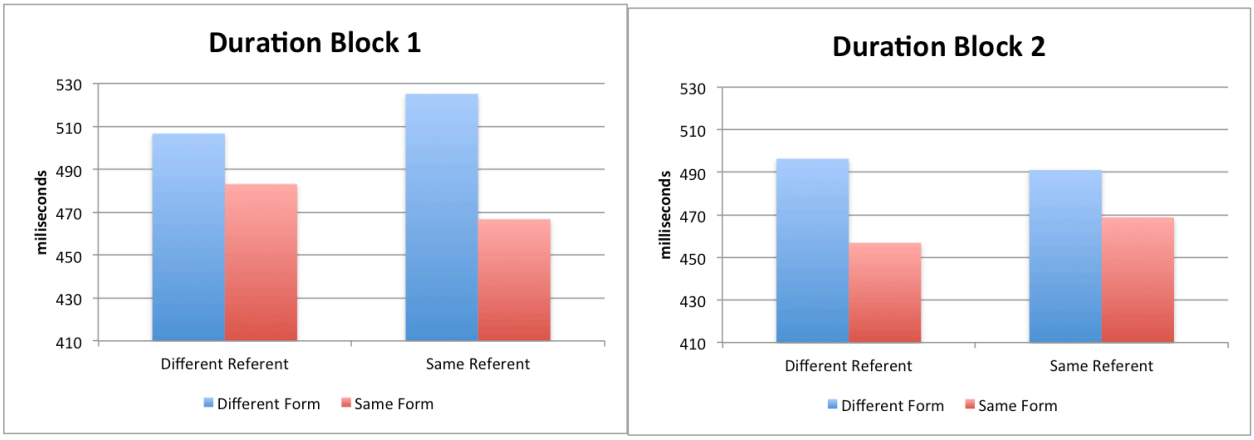

Figure 4.3. Shows the duration pattern in all four conditions across both across the first time speakers encounter the critical word in the experiment and the second time the speaker encounters the word.

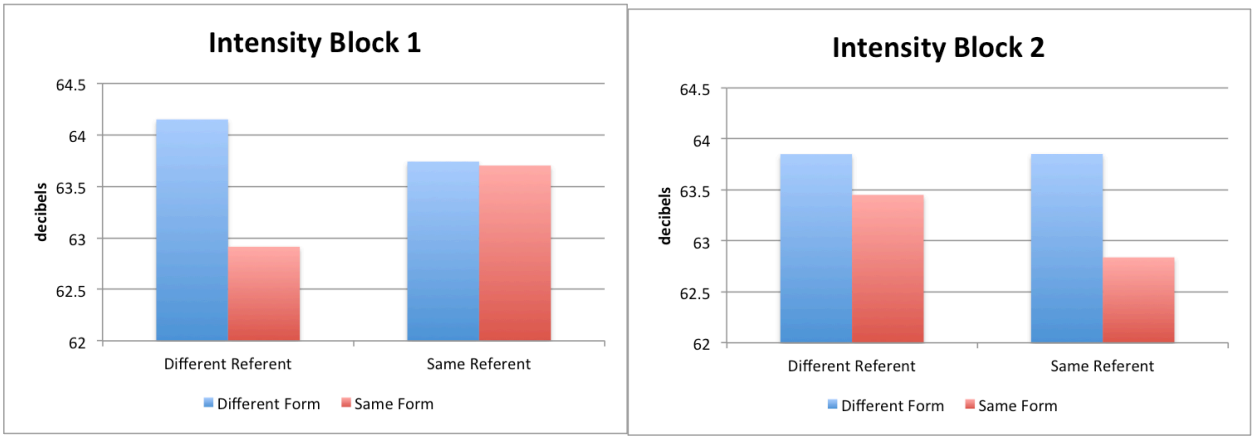

Figure 4.4. Shows the pattern for intensity for all four conditions across both across the first time speakers encounter the critical word in the experiment and the second time the speaker encounters the word.

Discussion

While the data pattern from this experiment is somewhat difficult to describe, there is one clear effect, which is that form repetition seems to matter most with respect to prominence reduction. For both intensity and duration, repeated forms led to reduced prominence. The results are much less clear for referent repetition. While referent repetition seems to affect prominence, it seems to interact with both form repetition and the number of times the word has been used in the experiment. Moreover, the pattern for duration differs from the pattern for intensity. For duration, repeating the referent with a different form led to longer durations when there was no repetition at all, but only on the first encounter. Perhaps part of this duration increase is due to contrastive stress due to changing the referring expression while maintaining the same referent. Recall that contrastive stress is when a word is produced with greater prominence in order to contrast it with a previously focused alternative (Pierrehumbert & Hirschberg, 1990). In the same referent, different name condition, there may be contrastive stress on the occupation name in order to contrast it with the previous mentioned occupation (i.e. previously the man was a soldier, but now he's a chef).

One problem with this explanation is that contrastive stress should also lead to greater intensity. However, this condition does not seem to stand out for intensity. Instead, the repeated referent, repeated form condition negates the form repetition effect such that repeated forms are produced with similar intensity as different forms in the repeated referent condition in the first encounter. This effect goes away by the second encounter such that the repeated referent, repeated form condition is produced with lower intensity than the repeated referent, different form condition. One possible explanation for this effect is that participants initially do not expect a given referent to change occupations at the start of the study and begin accommodating to this

occupational change after multiple trials. If this is true, one might expect that given enough trials, intensity would show an additive effect of referent and form repetition.

In conclusion, this experiment seems to support the idea that repetition reduction is driven primarily by form repetition. As previously discussed, this is likely not due to repetition of articulatory gestures but rather due to ease of retrieving a previously retrieved lexical item. Additionally, the experiment leaves open the possibility for some contribution from referent repetition at least for intensity. Overall, the pattern supports my general claim that prominence is affected by different levels of processing in production. However, the referent repetition effect seems relatively weak and may only operate after participants have come to treat the referent to be important in the task. To address the issue of whether speakers can learn to treat referent identity as an important factor in I conducted Experiment 7.

Experiment 7

Experiment 7 was a direct follow-up to Experiment 6, which found strong support for *form* repetition leading to reduced duration and intensity. However, the results for *referent* repetition were less clear. *Referent* repetition did seem to affect prominence but only for intensity and only in later trials. In Experiment 7, I attempted to replicate Experiment 6 but with more trials. If speakers accommodate to the occupation change over the course of the experiment, then they may begin to reduce prominence for repeated *referents* in later trials after they have become accustomed to this condition. By replicating Experiment 6 with more trials, it may be possible to examine prominence after speakers have become accustomed to the experimental conditions.

Method

Participants

Participants were a mix of subject pool participants from the University of Illinois and paid participants living in Champaign or Urbana, Illinois. Paid participants were compensated $8 for one hour of participation. Participants recruited from the subject pool were compensated with course credit in exchange for participation. All participants were native speakers of American English with normal or corrected to normal vision and hearing.

Thirty-three people participated in this study. Data from seven participants were excluded from analysis: Four participants failed to follow the instructions leading to unanalyzable data, one participant was excluded because of a sound equipment failure, one participant was not a native speaker of American English, and finally one participant had a cold and coughed during the experimental trials.

Materials

The materials were identical to the materials from Experiment 6, except with three replications of each occupation instead of two, leading to a total of 48 trials.

Procedure

The procedure was identical to the procedure from Experiment 6, except that participants in Experiment 7 completed 48 trials after the training phase instead of 32 trials.

Predictions

Most of the predictions are the same as in Experiment 6. If repetition reduction arises from repetition of the *referent* regardless of the referring expression used, then repeated *referents* should be reduced even when the occupation name has changed. If repetition reduction is due to repetition of the *form*, then repeated occupation names should lead to reduction, even when the referent has changed. Additionally, if speakers can accommodate to conditions in which the occupation changes, they may not initially reduce repeated *referents* upon the first encounter, but may begin reducing repeated *referents* by the second and third encounters. If this happens, it suggests that *referent* repetition reflects strategic control of prominence reduction whereas *form* repetition may be relatively automatic. Such a result would be consistent with a multiple source account of prominence.

Results

Prominence data were extracted using Praat, a speech analysis platform (Boersma & Weenink, 2007). Results were analyzed with multilevel linear regression with centered predictors. For both duration and intensity, the best fitting model included random intercepts for subject and items as well as a random slope term for block by subject suggesting that some subjects were more strongly affected by block than other subjects. All reported significant effects were also significant in an intercept only model. Data analysis for duration was conducted on raw durations whereas analysis for intensity was conducted on decibels.

There was a significant effect of lexical form repetition for duration such that repeated forms were produced with shorter duration than non-repeated forms ($t=8.641$, $p<0.0001$). However there was no significant effect of referent repetition on duration nor was there a

significant interaction with lexical repetition or with block; the pattern of results for duration was similar across all three blocks of the experiment. Figure 4.5 presents the overall pattern for duration data collapsed across the three blocks of the experiment.

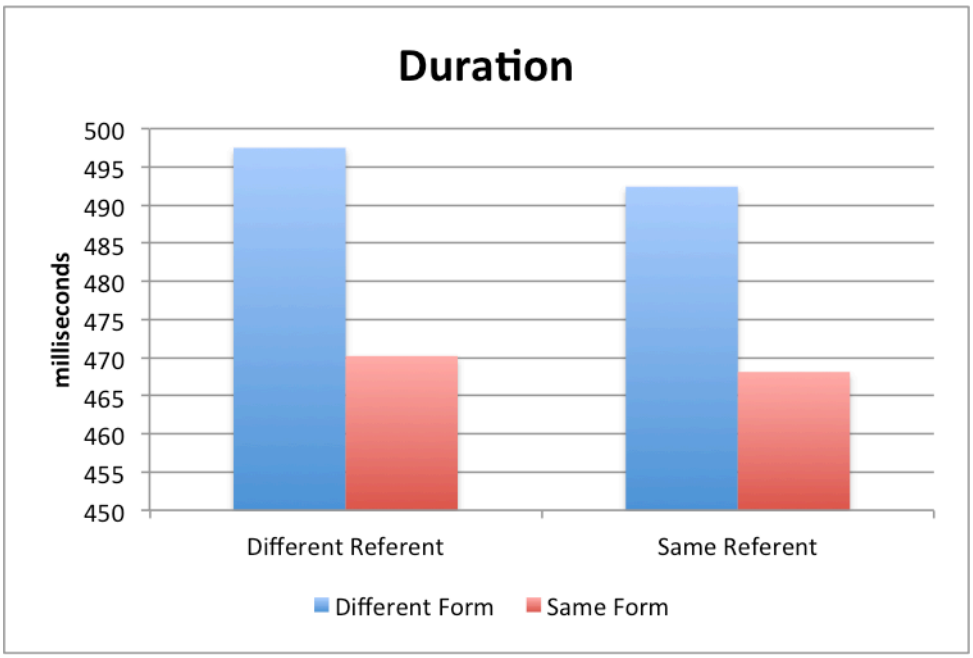

Figure 4.5. Depicts the overall pattern for duration collapsed across the three blocks of the experiment.

For intensity there was also a main effect of form repetition such that repeated forms were produced with lower intensity ($t=2.74$, $p<0.05$). Figure 4.6 presents the overall pattern for intensity collapsed across the three blocks of the experiment. While there was no significant main effect of referent repetition, there was a significant interaction between referent repetition and block such that speakers initially did increase intensity for repeated referents in the first block, but they reduce intensity in the second and third blocks ($t=2.31$, $p<0.05$). The pattern for intensity for the factor of referent repetition across the three blocks of the experiment is depicted in Figure 4.7.

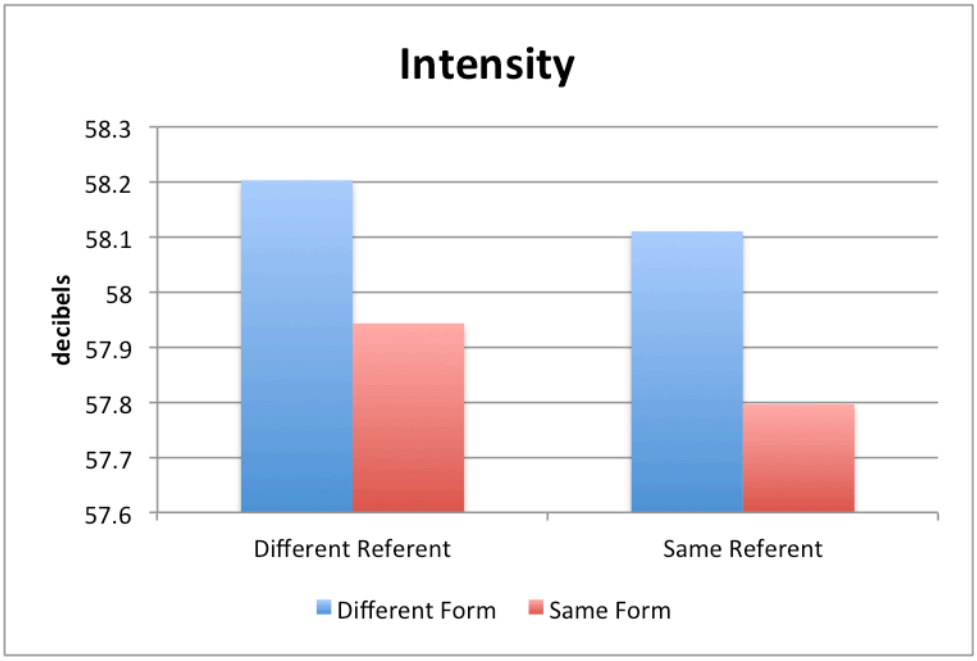

Figure 4.6. Depicts the overall pattern for intensity collapsed across the three blocks of the experiment.

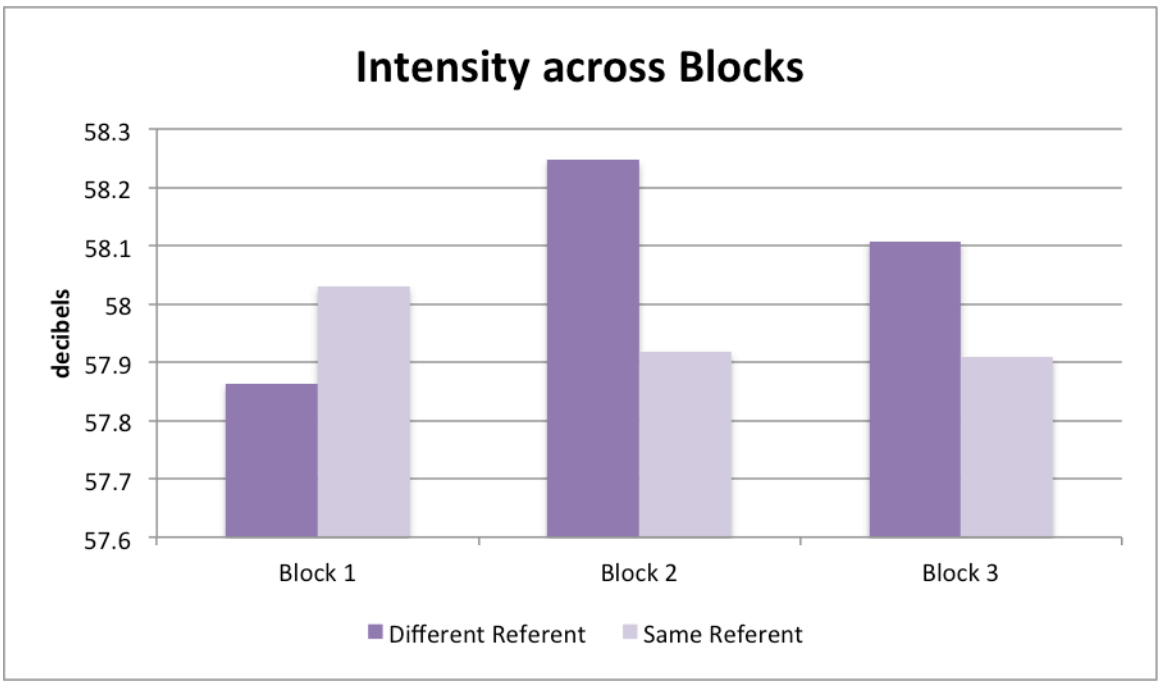

Figure 4.7. Depicts the pattern for referent repetition for intensity in each of the three blocks of the experiment. Data are collapsed across the factor of form repetition.

Discussion

Experiment 7 replicated the pattern from Experiment 6. There was a strong effect of lexical form repetition: repeated forms led to shorter durations and lower intensity regardless of whether the referent was repeated or not. Because repeated lexical forms led to reduction irrespective of whether the referent itself was repeated, it suggests that this effect is the result of processes at the level of lexical access, not at the message level. This pattern is most consistent with a lexical access theory of reduction. According to the lexical access theory of prominence reduction, words that have previously been mentioned are easier to retrieve (Bell et al., 2009; Lam & Watson, 2010). As a result of this ease of lexical access, the phonemes are retrieved more quickly leading to reduced prominence during production.

The data also suggest that the effect of repetition on reduction is not linked to message level planning. While referent repetition also seems to affect prominence, it does so only for intensity. Moreover, the effect only began in the second block of the experiment. This pattern replicates Experiment 6. One possible explanation for this effect is that participants did not expect a given referent to change occupations at the start of the study, and only accommodated to this occupational change after multiple trials. This effect may be similar to the predictability manipulation in Experiment 2 from Chapter, which showed that unexpected events at the referential level influence intensity more than duration in speech. The data from the current study is consistent with this interpretation. In Block 1, the condition with the greatest intensity numerically was the repeated referent, different occupation condition, and this condition may have been the most unexpected at the outset of the experiment.

General Discussion

Experiments 6 and 7 suggest that repetition reduction is primarily driven by lexical form repetition. In both experiments, reduction was linked to lexical repetition. This raises the question as to whether repetition of exact lexical forms is required for this sort of reduction or whether repetition of the phonetic form is sufficient for prominence reduction. Fowler (1988) argued that repetition of phonemic form is insufficient for reduction because homophone primes did not lead speakers to reduce duration. However, Gahl (2008) argues that so called homophones are not true homophones. The amount of duration reduction of a word like “time” and a word like “thyme” is predicted by the lemma frequency. That is, the amount of reduction is linked to lexical frequency, not to the frequency of the lexeme, which is the phonemic form (Gahl, 2008). If reduction is linked to the lemma frequency or the lemma to lexeme mapping,

then repeated lexical items should be reduced in prominence because repetition of the lexical item will lead to repetition of the lemma. This also predicts that words that are preceded by homophones that do not share the same lemma should not be reduced. Taken together with our study, this would suggest that the locus of the repetition effect is lexical repetition.

While Experiments 6 and 7 did find an effect of referent repetition, this effect only appeared in later trials. There are a couple of possible reasons why this effect is weak and appears only in later trials. As stated above, it is possible that speakers were initially surprised by the occupation change trials and actually produced repeated referents with greater intensity than non-repeated referents. In the second and third blocks speakers may have accommodated to these trials and then begin reducing intensity when referents are repeated. The fact that referent repetition affected only intensity may reveal something about how message level factors affect prominence. Both surprisal and referent repetition are factors that would operate at the level of the message. This pattern of results may suggest that effects that arise from factors at the message level will lead to prominence differences in intensity, at least for English. This interpretation is compatible with the predictability results from Experiment 2, which show reduction in intensity for predictable words in English. Because production planning begins at the level of the message and moves downward, it is possible that the intensity component of lexical repetition reduction may also arise from the level of the message.

In conclusion, Experiments 6 and 7 support the idea that repetition reduction is driven primarily by lexical form repetition. As previously discussed, this is likely not due to repetition of articulatory gestures but rather due to ease of retrieving a previously retrieved lexical item. Additionally, the experiment leaves open the possibility for some contribution from referent repetition, at least for intensity. However, the referent repetition effect was relatively weak and

may only operate after participants have come to treat the referent to be important in the task. Interestingly, repeated referents were produced with reduced intensity but not reduced duration whereas repeated form led to reduction in both duration and intensity. This pattern suggests a dissociation between duration and intensity in signaling different factors that affect prominence. Dissociation between duration and intensity are not without precedent (e.g. Experiment 2 in this dissertation; Watson, Arnold, & Tanenhaus, 2008).

Chapter 5: Conclusions

Multiple sources to prominence in the production system

As discussed in Chapters 1-4, intensity and duration do not always covary in signaling prominence. Previous researchers have argued that this may due to speech style or cue trading such that speakers signal prominence with one metric but not the other (e.g. Cole, Shattuck-Hufnagel, & Mo, 2009). However, the patterns described in the previous chapters instead suggest that message level factors and lexical level factors affect prominence in different ways. In English, these differences appeared primarily on intensity and duration. The pattern for English in Chapter 2 shows that intensity is more sensitive to task based predictability, whereas duration is more sensitive to prior mention by the speaker himself/herself. The pattern in Chapter 4, suggests that form repetition and referent repetition may be separable such that form repetition leads to reduction in both intensity and duration, but referent repetition may only lead to reduction in intensity. In Korean, repetition led to reduction in duration and F0, but predictability led increases in duration. These patterns are consistent with the multiple sources account of prominence such that lower-level processing-based factors lead to effects on prominence that are different from message based factors.

Recent work has also given some support to the idea that duration reflects low-level priming in the production system whereas intensity higher-level conceptual/message activation. For example, Kahn & Arnold (2010) found that givenness alone is not enough to explain effects of repetition reduction. In their study, speakers were asked to describe a set of objects from a computer display. They manipulated whether the objects in the display were linguistically given, conceptually given, or new. Linguistic givenness was manipulated by having the speakers first describe the objects in a prior event. Conceptual givenness was manipulated by cuing speakers to

which objects would be mentioned in the trial by having the objects flash on the screen. Utterances from these conditions were compared to utterances in which the objects were neither linguistically nor conceptual given. Kahn & Arnold found that both linguistically given and visually given referents were produced with shorter duration than new referents; however, linguistically given referents were more reduced than referents that were only conceptually given. From these results, they argued that conceptual givenness alone is not enough to explain repetition reduction. Rather, they argue that at least part of the repetition reduction effect is due to repetition in the production system itself.

In a separate study Fraundorf, Watson, & Benjamin (2010) showed that memory performance is predicted by reduction in speech intensity, but not duration. They asked speakers to complete a modified map task with a naïve listener. They elicited utterances of the form "Go from the noun1 to the noun2." On critical instructions, a previously mentioned (given) object is involved in the new instruction, and therefore is repeated. For all critical items, the authors computed the amount of reduction in intensity and duration from the first mention to the second mention. In both the first mention and the second mention, the object was in subject position. Following the map task, speakers were asked to recall all of the items that they had mentioned in the experiment. Items were then backsorted based upon whether or not they were recalled in the memory task. Items that were recalled showed less reduction in intensity than items that were not recalled. This was not true of duration. Recalled and forgotten items had similar levels of duration reduction. From this pattern, the authors concluded that reduction in intensity provides a cue to explicit recognition of words, whereas reduction in duration reflects low level priming effects in production.

The link between intensity and conscious awareness also applies in comprehension. In a comprehension study by Isaacs & Watson (2009), participants were asked to rate the prominence of different words. The words were extracted from natural speech and naturally varied in intensity, F0, and duration. The authors found that intensity was the best predictor of prominence ratings. That is, when listeners are consciously rating prominence of words, they rate according to intensity. In another study Kochanski, Grabe, Coleman, & Rosner (2005) built acoustic classifiers of prominence to determine which acoustic factors humans are most sensitive to in prominence judgments. These classifiers were trained on data from speech corpora annotated by trained phoneticians. They found that the classifier trained on intensity performed best overall. Moreover, while adding the other acoustic measures significantly improved classification, the improvement was small (76.6% vs. 78.6%).

The pattern of effects discussed above, suggests that duration and intensity are sensitive to different factors affecting prominence in English. Factors that influence ease of lexical selection for the speaker seem to affect duration (Chapter 2, Chapter 4, Bell et al., 2009; Kahn & Arnold, 2010, Galati & Brennan, 2010), while factors that affect the speaker's message (e.g. referential predictability, referent repetition, and addressee design) seem to affect intensity in English. However, this was not the pattern for Korean. In Korean, referential predictability and repetition both affected duration. However, even in Korean, the pattern suggests a difference between referential predictability and repetition. While the repetition effect in Korean is consistent with a processing based account of prominence reduction, the predictability effect cannot be easily accounted for based solely on an ease of processing account of prominence. This suggests that the effects come from two different sources despite the fact that they both affect duration. Overall, this pattern of results in the above chapters suggests that factors that

influence ease of lexical selection will also affect duration, and this effect seems to hold true across both English and Korean. Moreover, the fact that predictability affected English and Korean differently suggests that across different languages, message level factors may affect prominence in different ways.

The pattern of results in this dissertation suggests a dissociation between how message level factors and lexical level factors affect prominence. In Chapter 2, I showed that repeated words are reduced in duration simply because they are repeated. Because a word has previously been mentioned, subsequent repetition will lead to faster lexical retrieval and faster retrieval of the phonetic form. This should lead to shorter duration for repeated words. Even when repeated words are less predictable, they are still produced with shorter duration. Predictability and repetition affect different aspects of prominence. Moreover, the repetition effect holds in both English and Korean, a language that is argued to not even have pitch accents (Jun, 1993; 1998). In Korean, repeated words were produced with shorter duration while predictable words were produced with longer duration. Additionally, in Chapter 4, I showed that the repetition effect is due to repetition of *lexical forms*. Repeated lexical forms are produced with reduced duration and intensity. While repeated *referents* can also be produced with reduced prominence, this effect appears only in later trials and only on intensity. This pattern suggests that repetition at the level of the message affects prominence in a different way than that lexical repetition. Finally, while the results from Chapter 3 are somewhat ambiguous, the general pattern of results seems to suggest that addressee identity may affect prominence. Moreover, this effect seems to be sensitive to recent exposure to one's interlocutor because speakers who completed a cooperative verbal pretask with their addressees showed different prominence patterns from speakers who completed a silent, non-cooperative pretask. The fact that the pretask manipulation led to

reduced duration, but greater intensity and F0 may indicate that the manipulation affected multiple underlying factors that affect prominence. Perhaps duration was reduced due to ease of production from previously speaking to the addressee during the cooperative task, while intensity and F0 were increased because speakers wish to provide more informative productions to partners with whom they have recently worked with cooperatively.

Final remarks

The primary goal of this dissertation was to show that prominence effects arise from different levels of processing in the production system. In Chapter 2, I showed that repetition reduction is more than an effect of predictability. Repetition and predictability are separate factors that affect prominence. Repetition is primarily a lexical level effect while predictability is primarily a message level effect. In Chapter 3, I showed that the prominence of speakers' utterances is sensitive to the identity of their addressees, a message level effect. And in Chapter 4, I showed that the repetition effect itself, can be broken down into a message component and a lexical component, which affect prominence differently. Overall this pattern of results supports the idea that prominence not a binary variable, but is best thought of as a combination of multiple sources of prominence each with independent contributions to overall prominence.

References:

Andersen, A.H. & Howarth, B. (2002). Referential form and word duration in video-mediated and face-to-face dialogues. In J. Bos, M.E. Foster, and C. Matheson (eds.) Proceedings of the sixth workshop on the semantics and pragmatics of dialogue. (pp. 13-20). Edinburgh: University of Edinburgh.

Arnold, J. E. (1998). *Reference form and discourse patterns*. Unpublished doctoral dissertation, Stanford University.

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47, 31–56. doi:10.1177/00238309040470010201

Baker, R. E. and Bradlow, A. R. (2009) Variability in word duration as a function of probability, speech style and prosody. *Language & Speech*, 52(4), 391–413. doi:10.1177/0023830909336575

Bard, E.G., Anderson, A.H., Sotillo, C. Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000) Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*. 42(1), 1-22. doi:10.1006/jmla.1999.2667

Bard, E. G. & Aylett, M. P. (1999). The disassociation of deaccenting, givenness, and syntactic role in spontaneous speech. *Proceedings of the 1999 International Conference on Spoken Language Processing* (pp. 1753-1756).

Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge University Press.

Bell, A., Brenier, J., Gregory, M., Girand, C. and Jurafsky, D. (2009). Predictability Effects on Durations of Content and Function Words in Conversational English. *Journal of Memory and Language* 60:1, 92-111. doi:10.1016/j.jml.2008.06.003

Bell, A. Jurafsky D., Fosler-Lussier E., Girand C., Gregory M. L., & Gildea D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America* 113 (2), 1001-1024. doi:10.1121/1.1534836

Bock, K. (1995). Sentence Production: From mind to mouth. In J. L. Miller, & P. D. Eimas (Eds.), *Handbook of perception and cognition. Vol. 11: Speech, language, and communication* (pp. 181-216). Orlando, FL: Academic Press.

Bock, J.K. & Levelt, W. J. (1994) Language production. Grammatical encoding M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984), Academic Press, San Diego, CA.

Boersma, Paul & Weenink, David (2007). Praat: doing phonetics by computer (Version 4.5.14) [Computer program]. Retrieved from http://www.praat.org/

Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology*, 19, 441– 472.

Cole, J., Shattuck-Hufnagel, S., Mo, Y. 2010. Prosody production in spontaneous speech: Phonological encoding, phonetic variability, and the prosodic signature of individual speakers. *The Journal of the Acoustical Society of America*, 128: 2429.

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, 47, 292- 314. doi:10.1016/S0749-596X(02)00001-3

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, 5, 313-349. doi:10.1080/01690969008407066

Fowler, C. A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26, 489–504. doi:10.1016/0749-596X(87)90136-7

Fowler, C.A. (1988). Differential shortening of repeated context words produced in various communicative contexts. *Language and Speech, 31,* 307-319. Retrieved from http://las.sagepub.com

Fosler-Lussier, E., & Morgan, N. (1999). Effects of speaking rate and word predictability on conversational pronunciations. Speech Communication, 29, 137–158. doi:10.1016/S0167-6393(99)00035-7

Frank, A. and Jaeger, T.F. 2008. Speaking Rationally: Uniform Information Density as an Optimal Strategy for Language Production. *The 30th Annual Meeting of the Cognitive Science Society* (CogSci08), 933-938.

Fraundorf, S. H., Watson, D. G., & Benjamin, A. S. (2010, March). Recall predicted by reduction in intensity but not duration: Implications for theories of prominence. Poster presented at CUNY 2010: Conference on Human Sentence Processing, New York, NY.

Fussell, S. R. & Krauss, R. M. (1989). Understanding friends and strangers: The effects of audience design on message comprehension. *European Journal of Social Psychology*, 19, 509-525.

Galati, A. & Brennan, S. E. (2010). Attenuating repeated information: For the speaker, or for the addressee? *Journal of Memory and Language*, 62, 35-51.

Gregory, M.L., Raymond, W. D., Bell A., Fosler-Lussier E., & Jurafsky D. (1999). The effects of collocational strength and contextual predictability in lexical production. *In CLS-99* (pp. 151– 166). Chicago: University of Chicago.

Griffin, Z. M., & Bock, J. K. (1998) Constraint, word frequency, and relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, 38, 313-338. doi:10.1006/jmla.1997.2547

Gussenhoven, C. (1983). *A semantic analysis of the nuclear tones of English*. Bloomington (Indiana): Indiana University Linguistics Club.

Horton, W. S. (2007). The influence of partner-specific memory associations on language production: Evidence from picture naming. *Language and Cognitive Processes*, 22, 1114–1139.

Horton, W. S., & Gerrig, R. J. (2005). Conversational common ground and memory processes in language production. *Discourse Processes*, 40, 1–35.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59(1), 91–117. doi:10.1016/0010-0277(96)81418-1

Isaacs, A. M. & Watson, D. G. (2010). Accent detection is a slippery slope: Direction and rate of F0 change drives comprehension. *Language and Cognitive Processes*, 25, 1178-1200.

Isaacs, A. M. & Watson, D.G. (2009). Speakers and listeners don't agree: Audience design in the production and comprehension of acoustic prominence. Poster presentation at CUNY 2009: Conference on Human Sentence Processing, Davis, CA.

Isaacs, E. A., & Clark, H. H. (1987). References in conversations between experts and novices. *Journal of Experimental Psychology: General*, 116, 26-37.

Jaeger, T.F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*. doi:10.1016/j.cogpsych.2010.02.002

Jaeger, T. F. (2006). *Redundancy and Syntactic Reduction in Spontaneous Speech*. (doctoral dissertation), Stanford University.

Jescheniak, J. D., & Levelt, W.J.M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 824-843. doi:10.1037/0278-7393.20.4.824

Jun, S. A. (1993). *The Phonetics and Phonology of Korean Prosody*. Unpublished Ph.D. dissertation. The Ohio State University, Columbus, Ohio.

Jun, S. A. (1998). The accentual phrase in the Korean prosodic hierarchy, *Phonology*, 15(2), 189-226.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In Bybee, J., & Hopper, P. (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam: Benjamins.

Kahn, J. & Arnold, J.E. (2010, March). When predictability is not enough: the additional effect of givenness on acoustic reduction. Poster presented at the CUNY Sentence Processing Conference.

Kidd, C. & Jaeger, T.F. (2008, April) Prosodic Phrasing and Function Word Pronunciation. Spoken presentation given at Experimental and Theoretical Advances in Prosody, Cornell University, Ithaca, NY.

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005) Loudness predicts prominence: Fundamental frequency lends little. *Journal of the Acoustical Society of America*, 11(2), 1038–1054. doi:10.1121/1.1923349

Ladd, D. R. & Morton R. (1997). The perception of intonational emphasis: Continuous or categorical? *Journal of Phonetics*, 25, 313-342.

Lam, T. Q., Brown-Schmidt, S., & Watson, D. G. (2010, September). Speaker external and internal pressures on commonality assessment and audience design processes. Poster presented at Architectures and Mechanisms for Language Processing 2010, York, England.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–38.

Levy, R. & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In B. Schölkopf, J. Platt, and T. Hoffman (Eds.), *Advances in neural information processing systems* (NIPS) 19, 849-856. Cambridge, MA: MIT Press.

Lieberman, R. (1963). Some effects of the semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6, 172-175. doi:10.1121/1.1918465

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman.

Pierrehumbert, J. (1980) The Phonology and Phonetics of English Intonation. Ph.D thesis, MIT.

Pierrehumbert, J. and J. Hirschberg (1990) The Meaning of Intonational Contours in the Interpretation of Discourse, in P. Cohen, J. Morgan, and M. Pollack, (Eds). *Intentions in Communication*, MIT Press, Cambridge MA. 271-311.

Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005a). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica*, 62, 146-159. doi:10.1159/000090095

Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005b). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America*, 118, 2561-2569. doi:10.1121/1.2011150

Rossion, B. & Pourtois, G. (2001). Revisiting Snodgrass and Vanderwart's object database: Color and texture improve object recognition. *Journal of Vision*, 1(3), 413a.

Savitsky, K., Keysar, B., Epley, N., Carter, T. and Swanson, A. (2011). The closeness-communication bias: Increased egocentrism among friends versus strangers. *Journal of Experimental Social Psychology*, 47, 269-273.

Schwarzschild, R. (1999). Givenness, AVOIDF and other constraints on the placement of accent. *Natural Language Semantics*, 7, 141-177.

Selkirk, E. (1995), Sentence Prosody: Intonation, Stress, and Phrasing, In: J. A. Goldsmith (ed.): *The Handbook of Phonological Theory*. London: Basil Blackwell, pp. 550–569.

Shannon, C. E. (1951). Prediction and entropy of printed English. Bell System Technical Journal, 30, 50–64.

Snodgrass, J. G. & M. Vanderwart (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory* 6(2), 174-215. doi:10.1037/0278-7393.6.2.174

Watson, D.G. (2010). The many roads to prominence: Understanding emphasis in conversation. In B. Ross (Ed.) *The Psychology of Learning and Motivation* (pp.163-183), Vol. 52. Burlington: Academic Press.

Watson, D. G., Arnold, J. E. & Tanenhaus, M. K. (2008) Tic Tac TOE: Effects of predictability and importance on acoustic prominence in language production. *Cognition*, 106, 1548-1557. doi:10.1016/j.cognition.2007.06.009

Wells, H. G. (1898) *The War of the Worlds*. Bartleby.com, 2000. www.bartleby.com/1002/.

Wingfield, A. (1968). Effect of frequency on identification and naming objects. *American Journal of Psychology*, 81, 226-234. doi:10.2307/1421267

Zipf, G. K. (1929). Relative frequency as a determinant of phonetic change. *Harvard Studies in Classical Philology*, 15, 1-95.